DOCUMENT RESUME

ED 077 268 FL 003 828

AUTHOR	Grosu, Alexander
TITLE	The Strategic Content of Island Constraints. Working Papers in Linguistics No. 13.
INSTITUTION	Ohio State Univ., Columbus. Dept. of Linguistics.
PUB DATE	Dec 72
NOTE	229p.; Modified version of Ph.D. dissertation, Ohio State University, August 1972
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	*Behavioral Sciences; *Descriptive Linguistics; Doctoral Theses; English; Grammar; Language Research; Language Universals; Linguistic Competence; Linguistic Theory; *Perception; *Psycholinguistics; Semantics; Surface Structure; *Syntax; Transformation Generative Grammar; Transformations (Language)

ABSTRACT

This dissertation seeks to provide an improved syntactic analysis of the phenomena known as "island constraints" and to offer a natural explanation for these and other phenomena in behavioral terms. The author approaches the problem using the notions of "nucleus" and "satellite," and shows that these concepts are a necessary part of a descriptively adequate grammatical theory. The behavioral principles seen to be at work are erroneous closure, interrupted behavior, and perceptual conflict. The author presents a reanalysis of three constraints (the complex noun-phrase constraint, the coordinate structure constraint, and a third constraint involving adverbials) and uses the behavioral principles listed above to propose explanations for the main island constraints under consideration. (Author/VM)

ED 077268

WORKING PAPERS IN LINGUISTICS NO. 13

THE STRATEGIC CONTENT OF ISLAND CONSTRAINTS

By

Alexander Grosu

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Department of Linguistics
The Ohio State University
Columbus, Ohio 43210

December, 1972

FL 003 828

PREFACE

This issue of Working Papers in Linguistics is a slightly modified version of Alexander Grosu's Ph.D. dissertation from The Ohio State University (submitted to the Graduate School in August, 1972).

Mr. Grosu's address is: Department of English Linguistics
Tel-Aviv University
Ramat-Aviv, Israel

Copies of Working Papers in Linguistics No. 13 are available from the Department of Linguistics at Ohio State University for $2.00.

ACKNOWLEDGMENTS

I wish to express my deepest gratitude to my adviser, Professor Arnold Zwicky, who has read at least five versions of this dissertation and who, through countless invaluable comments, has in fact made it possible; to Professors Ilse Lehiste, Michael Geis, Gaberell Drachman, and Francis Utley, who have offered many pertinent helpful suggestions; to the Graduate School of The Ohio State University, which awarded me a Dissertation Year Fellowship; to Professor Robert Lees and to Tel-Aviv University, for their encouragement of every kind since the day I became a freshman in Linguistics; to my informants, Miss Mieko Ohso and Mr. Ahmad Siddiqui, who have provided the Japanese and Hindi data respectively; and, last but not least, to my wife, Mariana Grosu, who has put up with the strain of a dissertation year while offering her constant encouragement.

TABLE OF CONTENTS

Page

PREFACE ii

ACKNOWLEDGMENTS iii

INTRODUCTION 1

Chapter

I. THE TRANSFORMATIONAL AND DERIVATIONAL POSITIONS . . . 6

II. PERCEPTUAL PRINCIPLES AND GRAMMAR 50

III. THE 'NUCLEUS-AND-SATELLITE' CONSTRUCTION 110

IV. INTERRUPTION, CONFLICT, CLOSURE, AND ISLAND CONSTRAINTS 153

V. SUMMARY AND FURTHER OUTLOOKS 206

APPENDIX

1. 208

2. 212

3. 216

4. 219

BIBLIOGRAPHY 222

INTRODUCTION

This dissertation constitutes an attempt to provide explanations for the syntactic phenomena currently known as "Ross' constraints."

While there have been a number of attempts aimed at providing more adequate (in terms of generality, simplicity, or some such criterion) reformulations of the phenomena in question, I know of no proposal which seriously addresses itself to the task of explaining them. To be more exact, most discussions of Ross' constraints have regarded them as formal syntactic universals, and have assumed that an adequate statement of them would constitute their ultimate explanation.

The position adopted in this dissertation is that a mere statement of the constraints at issue, even if observationally and/or descriptively adequate (which most treatments with which I am familiar are not, as will be shown below), should not be regarded as an explanation of the facts. The point of view expressed in the preceding sentence applies with equal force to the position that Ross' constraints constrain transformations (henceforth: the Transformational Position)--taken, for example, in Chomsky (1964a, 1964b, and 1971), and Ross (1967, and an unpublished proposal known as the Island Constraint)--and to the position that the constraints in question are global restrictions on derivations (henceforth: the Derivational Position)--as taken, for example, in Lakoff (1969), Postal (1969), and Ross (1969b). The reason why I have serious doubts that a mere transformational or derivational statement can constitute the explanation of the constraints in question is that it is unlikely, given the present state of the field, that any version of transformational grammar known at present could be an adequate account of the competence of language users. The large number of counterexamples which have so far plagued any attempt to provide a reasonably elegant account of significant bodies of data, as well as the results of recent mathematical investigations of the properties of transformational grammars (e.g., Peters and Ritchie, 1968), strongly suggest that a transformational grammar is not a realistic model of linguistic competence. If so, it is not in the least clear what, if anything, corresponds to the notion "transformation" in psychological terms, and claims to the effect that formal properties of these doubtful entities--the transformations--constitute explanations of observable facts must be taken with a considerable pinch of salt.

Given the questionable explanatory status of purely formal universal constraints, there remain two classes of facts which may provide explanations for syntactic data: semantic and behavioral ones. Such facts are on considerably surer grounds, for, while we may doubt the existence of transformations, there are no reasons for doubting the existence of concepts, or of perception and production mechanisms.

With respect to Ross' constraints, there are clear indications that they have very little (if anything) to do with semantics (in fact, Ross himself repeatedly emphasized that his constraints concerned derived, rather than underlying, structure). Thus, consider the fact that sentences with sentential subjects are subject to one of Ross' constraints, while the transformational congeners of such sentences in which the sentential subject is extraposed are free from the pertinent constraint, even though the two sentence-types have identical underlying representations; consider also the fact that in an analysis which regards relative clauses as derived from conjuncts, the clause containing the relative is free from constraints, while the conjunct which underlies it is not; notice also that certain rules may operate either by moving or by copying a constituent, and, although the two kinds of processes yield exact paraphrases, only the former is subject to Ross' constraints; notice, finally, that certain feature-changing processes, such as the one that places an overt mark on question-phrases, is free from constraints in Japanese, but not in English. I believe that the few facts mentioned above show quite conclusively that Ross' constraints cannot be explained in terms of semantics, for, if semantic representations were subject to such constraints, any surface structure derived from ill-formed representations should itself be ill-formed, which, as we have seen above, is not the case.

As a semantic explanation is not available, and as a mere formal statement cannot be regarded as an explanation (at least at the present time), it remains to investigate the possibility that sentences in which Ross' constraints have been violated conflict with certain properties and/or limitations of the perceptual apparatus; the main goal of this dissertation is to investigate this possibility. In doing this, I will rely, to some extent, on the theory of the interaction of perception and grammar proposed in a number of papers by Fodor, Garrett, Bever, Langendoen, and others. However, given the highly tentative and often questionable status of many of its substantive claims, that theory will be used as sparingly as possible. Specifically, I will confine myself to three types of perceptual principles in proposing explanations for syntactic phenomena, namely, principles which invoke the complexity resulting from erroneous closure, interrupted behavior, and perceptual conflict respectively. Closure principles assert that complexity arises when a proper subpart of a structure is mistakenly apprehended as a well-formed subpart of that structure, with the result that the remainder of that structure appears ill-formed; interruption principles assert that a proper subset of the set of discontinuities create complexities in proportion to the values of specific parameters, such as the length, structural complexity, overall structural predictability, etc., of the intervening material; conflict principles assert that complexity arises when two sets of cues assign contradictory values to a stimulus in terms of some parameter.

In using closure, interruption, and conflict principles, I shall attempt to narrow down each principle as much as possible, for merely saying that erroneous closure, interrupted behavior, or perceptual conflict are complex is general to the point of near-vacuity. Each proposed principle will be constrained to the extent allowed by the data, considerations of plausibility, and past experimental findings, but no attempt will be made to offer a definition of the notion "possible perceptual principle"; it seems reasonable to assume that such a definition should be hard, if not impossible, to come by in the absence of an adequate and comprehensive theory of linguistic perception, an extremely remote goal at present. In fact, I believe that a definition of the notion "possible perceptual principle" should be a major *goal* of psycholinguistic research, rather than one of its *prerequisites*. In this sense, each proposed principle may be viewed as a *tentative partial* definition of the notion "possible perceptual principle". Clearly, the tentative character of the proposals I shall make is undeniable, and follows directly from the non-existence of a satisfactory theory of linguistic perception; the situation is, however, no worse than that obtaining in other fields of linguistic investigation, for no one has, to the best of my knowledge, proposed an adequate definition of the notions "possible transformation", "possible global constraint", "possible phonological rule", etc., so far.

The relation between behavioral complexity and acceptability judgments is fairly transparent in some of the cases I consider; for example, it is rather easy to show that the relative acceptability of discontinuities varies as the length and/or structural complexity of the intervening material is varied. However, not all situations are equally clear in this respect, as (a) some parameters, especially in cases involving conflict, often exhibit a very limited number of possible values, and (b) even when an arbitrarily large number of values is possible with respect to some parameter, the correlation between perceptual complexity and degree of acceptability is sometimes obscured through *generalization*, a process which imposes a yes/no dichotomy on the scalar acceptability dimension; the underlying[1] correlation can in fact be further obscured by reinterpretation, generalization, and other extremely common phenomena in language change. Such difficulties should not, however, preclude inquiries into the possibility that certain ungrammatical constructions acquire this status *because of* some high degree of perceptual complexity, just as the fact that certain rule-systems are synchronically unproductive has not, in general, precluded research as to the possibility of an earlier productive state. In both situations, empirical claims are made, and it should be possible to test them, in the latter case through an examination of historical records, in the former, through psycholinguistic experimentation (of course, tests aimed at proving a putative underlying perceptual-complexity/acceptability correlation can obviously not be applied directly to structures which have undergone grammatization, but rather to (a) the same structures, in dialects (or languages) where grammatization has not taken place, (b) other structures, in the same dialect, provided that the relevance of these structures to the primary ones can be defended, or (c) representations in other perceptual modalities, subject to the condition expressed with respect to (b)).

---

[1] I am using the term *underlying* rather than *initial*, to allow for the possibility that some instances of grammatization of perceptual restrictions do not necessarily occur at some stage in the historical evolution of a language, but more or less inevitably for each individual learner (presumably due to an intolerably high degree of complexity). Therefore, I am *not* making the claim that a sentence like **he$_i$ said that John$_i$ had left* was necessarily grammatical at some point in the history of English.

It would be extremely interesting to try to find out whether there is some quantifiable threshold of perceptual complexity which when reached, allows reasonably certain predictions that grammatization will occur. To ask this question meaningfully, it would be necessary, at the very least, to know (a) what all the perceptual variables which may increase or reduce complexity are, and (b) what weights should be assigned to the various variables in situations where a subset (not necessarily proper) of them interact. As the answer to the former, let alone the latter, question is not known at the present time, it is clear that any attempt to compute such a grammatization threshold would be hopelessly premature at this stage. For example, one may attempt the generalization that "total", as distinct from "partial", perceptual conflict is necessarily grammatized (for a discussion of the terms in inverted commas, see section 4.3); however, this "generalization" will be seen to be incorrect in general, precisely because there exist mitigating, in addition to aggravating, factors in perception.

This dissertation is incomplete in two important ways: (a) its theoretical claims are seriously underdetermined by the data, as I have not been able to gather enough pertinent facts from a sufficiently rich sample of languages, and (b) while it is consistent with the results of past experiments, it is not supported by any especially designed experiments, owing to the fact that I have had neither the time nor the facilities for devising and carrying out pertinent tests. Fortunately, both shortcomings can be remedied through future work. I believe that despite these two objections whose seriousness should not be underestimated, this dissertation does make a contribution to the study of syntax by providing an account which is empirically and descriptively superior, as well as more elegant, plausible, and intuitively satisfactory, than previous treatments of the same problems.

The breakdown of this dissertation is as follows:

Chapter One examines the most important earlier proposals made within the Transformational Position and shows them to be inadequate on both empirical and explanatory grounds. The Derivational Position, which has never been spelled out in detail, is pursued to some extent, and it is argued that empirical adequacy could be achieved at the cost of making unrestricted use of global constraints, as well as allowing reference to degrees of acceptability and to pragmatic information; it is further argued that explanatory adequacy would be highly improbable under those circumstances, since it is hard to see what could not be described with such a powerful apparatus.

Chapter Two consists of three parts: section 2.0 briefly outlines the theory of perceptual strategies sketched in Bever (1970); section 2.1 provides illustrations of the applicability of the closure, conflict, and interruption, principles, and focuses essentially on cases in which the complexity/acceptability correlation is relatively transparent; section 2.2 examines the validity of a putative perceptual strategy (whose correctness has usually been assumed) to the effect that syntactic choices are made on the basis of "minimal distance" considerations (the plausibility of this assumption being probably due to the independently attested "recency effect" in recall experiments); it is argued that there is little or no evidence that minimal distance principles play a significant part in syntax; such principles are often in conflict with the facts, and, when not in conflict, more convincing alternative explanations are available.

Chapter Three presents a detailed defense of my proposed reanalysis of three constraints (the Complex NP Constraint, the Coordinate Structure Constraint, and a third constraint involving adverbials), which, I argue, fall under the same generalization.

Chapter Four employs the three perceptual principles discussed in Chapter Two to propose explanations for the main island constraints in the light of the reanalysis of Chapter Three.

Chapter Five summarizes the results of the dissertation.

CHAPTER I

THE TRANSFORMATIONAL AND DERIVATIONAL POSITIONS

1.0. The main subdivisions of this chapter, namely, 1.1, 1.2, 1.3, and 1.4, are respectively concerned with Chomsky's initial proposal known as the A-over-A Principle, with the extensive study made in Ross (1967), with Chomsky's reanalysis of much the same facts within the framework of his Extended Standard Theory, and with Ross' attempt to collapse most of his constraints into an Island Constraint; all these proposals espouse the Transformational Position. The Derivational Position is discussed in 1.2.6 in relation to Ross' original position.

1.1.1. In Chomsky (1964a), it was pointed out that the question and relative clause transformations are ambiguous when applying to a noun phrase modified by a full or reduced relative clause. Thus, in (1.1), both the phrase the boy from Los Angeles and its head the boy are noun phrases; however, the question or relative clause transformations must apply only to the former, yielding the sentences in (1.2), rather than to the latter, as this would yield the ungrammatical sentences in (1.3).

(1.1) You believe the boy from Los Angeles to be unbalanced.

(1.2) a. Who do you believe to be unbalanced?
b. The boy from Los Angeles who you believe to be unbalanced has won the Nobel Prize.

(1.3) a. *Who do you believe from Los Angeles to be unbalanced?
b. *The boy who you believe from Los Angeles to be unbalanced has won the Nobel Prize.

Chomsky assumes that transformations must be unambiguous and proposes a hypothetical linguistic universal that will eliminate the ambiguity. This universal (which, following Ross, I shall call the A-over-A principle) was stated formally as follows:

(1.4) ...if the phrase X of category A is embedded within a larger phrase ZXW which is also of category A, then no rule applying to the category A applies to X (but only to ZXW).

1.1.2. In Chomsky (1964b), a revised version of Chomsky (1964a), it is pointed out in note 10 that the A-over-A principle is too strong, as it would predict that sentences like (1.5) are ungrammatical.

(1.5) Who do you approve of my seeing?

(1.6) You approve of my seeing someone.

Indeed, on the assumption that (1.6) is the source of (1.5), it contains the NP someone embedded within the larger NP my seeing someone, and still the question transformation can apply to the lower NP without ungrammaticality.

I should like to point out at this stage that the failure of the A-over-A principle to allow the generation of (1.5) depends crucially on the overall structure of the grammar which incorporates the principle. Specifically, it depends on my seeing someone being dominated by the node NP. If the A-over-A principle is incorporated into a grammar which generates all complement sentences under the domination of the node NP (as proposed in Stockwell et al., 1968, for example), then it becomes a trivial matter to show that the A-over-A principle is incorrect, for it would predict that no NP moves out of an embedded clause, and this is clearly not the case (at least for English).

1.2.1. In Ross (1967), the most thorough treatment of the constraints at issue within the Transformational Position, Ross mentions Chomsky's observation with respect to sentences like (1.5), and adds that the A-over-A principle is also too strong with respect to certain cases which exhibit an unbounded sequence of NPs, such that for any two NPs X and Y, X either dominates Y or is dominated by it. Specifically, given a string like (1.7), the A-over-A principle predicts that (1.8a) and (1.9a) alone are grammatical, and incorrectly rules out (1.8b, c) and (1.9b, c).

(1.7) You saw a picture of the master of the house.

(1.8) a. What did you see?
b. Who did you see a picture of?
c. What did you see a picture of the master of?

(1.9) a. It's a picture of the master of the house that you saw.
b. It's the master of the house that you saw a picture of.
c. It's the house that you saw a picture of the master of.

The failure of the A-over-A principle in cases like (1.7), where nodes of some type A can be recursively embedded to the right of other nodes of Type A, is in fact more general, since the variable A can stand for S or VP, not just NP. Thus, Topicalization and Pseudo-Clefting can apply to any of the S nodes in the underscored portion of (1.10), yielding (1.11) and (1.12) respectively, while Pseudo-Clefting can apply to any of the VP nodes in the underscored portion of (1.13), yielding (1.14).

(1.10) I believe that John claimed that Mary insinuated that Bill was guilty of murder.

(1.11) a. John claimed that Mary insinuated that Bill was guilty of murder, I believe.
b. Mary insinuated that Bill was guilty of murder, I believe that John claimed.
c. Bill was guilty of murder, I believe that John claimed that Mary insinuated.

(1.12) a. What I believe is that John claimed that Mary insinuated that Bill was guilty of murder.
b. What I believe that John claimed is that Mary insinuated that Bill was guilty of murder.
c. What I believe that John claimed that Mary insinuated is that Bill was guilty of murder.

(1.13) John decided to try to begin to write a book.

(1.14) a. What John decided was to try to begin to write a book.
b. What John decided to try was to begin to write a book.
c. What John decided to try to begin was to write a book.

1.2.2. After having considered the cases which were incorrectly ruled out by the A-over-A principle, Ross turns, in section 2.2, to the six[1] cases in (1.15) which, according to him, can be handled by that principle in a satisfactory manner:

(1.15) a. Elements of relative clauses cannot be questioned or relativized; thus (1.16) is ungrammatical.

(1.16) *This is the girl who I know a boy who likes.

(1.15) b. Elements of sentences in apposition to "sentential" nouns like fact, idea, question, etc., may not be questioned or relativized, as may be seen in (1.17):

---

[1] Ross also mentions a seventh case, suggested by McCawley and involving the Adjective Shift Rule, but concludes in section 2.3 that the rule itself is inadequate and that a more satisfactory formulation would no longer require the A-over-A principle.

(1.17) *Who did I mention to you the fact that John seduced?

(1.15) c. A clause modifying a head NP cannot be extraposed beyond "the first sentence up", as shown in (1.18).

(1.18) *A proof that the claim had been made was given that John had lied.

(1.15) d. The head of a relative clause cannot be questioned or relativized, and neither can a noun like fact or idea when it is modified by a clause (Ross omits to mention the latter case); these two restrictions are exhibited in (1.19a, b) respectively:

(1.19) a. *Who did he expect who I was acquainted with would show up?
b. *What do I believe (that) that John had lied is well established?

(1.15) e. A NP which is exhaustively dominated by a Determiner cannot be moved out of the NP which immediately dominates that Determiner, as in (1.20):

(1.20) *Whose did you buy house?

(1.15) f. A conjunct or a disjunct NP in a coordinate node cannot be moved out of the latter, as seen in (1.21). In fact, the A-over-A principle prevents the movement of a member of any coordinate node, and this restriction is necessary, for Ss and VPs can also move, as seen in (1.11), (1.12), and (1.14), and the ungrammatical (1.22) and (1.23) must be prevented.

(1.21) a. *What will you put between the bed and?
b. *What will you put between and the wall?

(1.22) a. *What John claimed that Mary left and was that Bill ran away.
b. *What John claimed and that Bill ran away was that Mary left.

(1.23) a. *What John likes to eat and is to drink.
b. *What John likes and to drink is to eat.

1.2.3. Having shown that the A-over-A principle was too strong, Ross proposes to handle the cases which the principle was adequate for by four separate constraints. These four constraints taken together are weaker than the A-over-A principle, as they account for the (1.15a)-(1.15f) cases without starring the sentences in (1.5), (1.8), (1.9), (1.11), (1.12) and (1.14), but they are also stronger, as they can account for additional cases. This last fact does not constitute a defect of the A-over-A principle, for the latter was never meant to account for all the constraints in a grammar; it does, however, strengthen Ross' analysis. The main way in which Ross' constraints strengthen the A-over-A principle concerns the fact that elements of conjuncts or disjuncts cannot move out of the coordinate node; this constraint is not expressed by the A-over-A principle in those cases where the coordinate node and the elements prevented from moving belong to different categories, as in (1.24a), where the underscored element cannot move to yield (1.24b):

(1.24) a. John loves Mary and Bob hates Jill.
b. *It's Mary who John loves and Bob hates Jill.

Sentences like (1.25b) and (1.26b), whose ungrammaticality is due to the migration of the nodes VP and ADJ respectively from a complex NP, could also be claimed to strengthen Ross' reformulation of the A-over-A principle.

(1.25) a. I believe the claim that John decided to try to write a book.
b. What I believe (*the claim) that John decided to try was to write a book.

(1.26) a. John heard the claim that Mary was pretty.
b. Pretty though John heard (*the claim) that Mary was, he never showed much interest in her.

However, these sentences are not relevant in the model of grammar Ross proposes, for he regards VP's and predicative ADJ's as dominated by the category NP.

1.2.4. The four constraints Ross puts forward in an attempt to overcome the inadequacies of the A-over-A principle are: The Complex Noun Phrase Constraint, The Coordinate Structure Constraint, The Left Branch Condition on the Pied-Piping Convention, and The Sentential Subject Constraint (henceforth, the CNPC, the CSC, the LBC on the PPC, and the SSC respectively). The first two are hypothesized to be universals,[2] the last two are claimed to be language-specific.

---

[2] The universality hypothesis was subsequently shown to be false by a number of investigators. Thus it appears that the Complex NP Constraint does not hold in Swedish (D. Perlmutter, personal communication) or in Portuguese (C. Quicoli, personal communication). This demonstrates the incorrectness of various attempts to explain Ross' constraints on semantic grounds, since, if the structures at issue were ill-formed semantically, there should

be no exceptions in any language whatever. In contrast, the explanation which I shall propose, and which ultimately involves the violation of conversational and/or perceptual principles is not defeated by a few counterexamples, since principles of this kind can be violated.

(1.27) The CNPC: No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation.

Ross treats the feature [±Lexical] as a formal one, but the examples he gives from English and Japanese suggest that the plus-value is associated with semantically empty forms, like the English it that results from Extraposition, or the Japanese koto, mono, which translate roughly as "thing".

The CNPC takes care of cases (1.15a) and (1.15b) that were accounted for by the A-over-A principle, and can also block the movement of any kind of element out of the modifying clause, while the A-over-A principle can only block NPs.

(1.28) The CSC: In a coordinate structure, no conjunct may be moved, nor may any element contained in that conjunct be moved out of that conjunct.

It is clear that the first part of this constraint has the same effect as case (1.15f) of the A-over-A principle, while the second part covers the migration of elements of coordinate terms, not covered by the latter. A number of counterexamples to (1.28) (involving asymmetric coordination and across-the-board operations) which Ross brings up will be discussed in detail in Chapter Three.

(1.29) The PPC: Any transformation which is stated in such a way as to effect the reordering of some specific node NP, where this node is preceded and followed by variables in the structural index of the rule, may apply to this NP or to any non-coordinate NP which dominates it, as long as there are no occurrences of any coordinate node, nor of the node S, on the branch connecting the higher node and the specified node.

(1.29) has the effect of allowing the sentences (1.8b, c) and (1.9b, c) which the A-over-A principle ruled out. As I have already pointed out (1.29) needs to be strengthened in order to account for the b and c cases in (1.11), (1.12) and (1.14). Of the five language-specific conditions that Ross imposes on the PPC, I shall mention only the LBC as the other four strike me as truly idiosyncratic grammatical facts and thus of little interest for the Behavioral Position I have espoused.

(1.30) The LBC: No NP which is the leftmost constituent of a larger NP can be reordered out of this NP by a transformational rule.

The LBC accounts for cases (1.15d) and (1.15e) of the A-over-A principle.

The fourth constraint Ross proposes, the SSC, accounts for the unacceptability of (1.31b) in contrast to the acceptability of (1.32b) and (1.33b).

(1.31) a. That John loves Mary is odd.
b. *It's Mary who that John loves is odd.

(1.32) a. It's odd that John loves Mary.
b. It's Mary who it's odd that John loves.

(1.33) a. Bill claims that John loves Mary.
b. It's Mary who Bill claims that John loves.

(1.34) The SSC: No element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S.

It is not clear whether the A-over-A principle was initially intended to account for sentences like (1.31b), for it is not clear whether Chomsky considered complement sentences to be dominated by NP. I would imagine that he did not, for counterexamples like (1.32b) and (1.33b) are much too obvious. In any event, in a grammar which does represent all complement sentences as dominated by NP, the A-over-A principle is much too strong, and the SSC is a welcome weakening.

The four constraints mentioned so far do indeed improve on the A-over-A principle. However, there is another class of configurations which impose constraints on movement transformations, namely, the adverbials, and neither the A-over-A Principle nor the four above-mentioned constraints handle these cases adequately. Thus, it seems that at least the optional (in a sense to be made more precise in Chapter Three) adverbials cannot lose elements. In talking of adverbials, I shall refer ambiguously to the entire adverbial, or to the adverbial minus the "connector" which introduces it (preposition, 'subordinating conjunction', or whatever; the term 'connector' is used for purely mnemonic purposes, and no claim is made that a category Connector should be part of the non-terminal vocabulary of the grammar); the reason is that the 'smaller' adverbial as a whole can sometimes be moved (even though it is an element of the 'larger' adverbial), but elements of the smaller adverbial cannot be moved.

In Ross (1967), the adverbs are treated as a subset of the complex NP category. That is, it is assumed that all adverbs are complex NP's at some level of representation, and that movements out of adverbial clauses occur prior to some transformation which substitutes subordinating conjunctions for the heads of complex NP's (and, perhaps, some more prelexical material). The hypothesis that adverbs originate as complex NP's is due to M. Geis, who supports it extensively, and, I think, convincingly, in Geis (1970). Even though Geis does not carry out a detailed analysis of all the forms which traditional grammarians have called "subordinating conjunctions", he strongly implies that all adverbs should be analyzed as complex NP's.

Given the strong evidence supplied by Geis for the cases he analyzed, there is no reason to reject outright the hypothesis that all subordinating conjunctions may originate as complex NP's. However, even if such an analysis ultimately turns out to be correct, there are several cases involving adverbials in which movements are blocked and which it cannot account for.

One such case concerns the parentheticals, which, being (sentential) modifiers are (at least surface) adverbials. As (1.35) shows, such parentheticals become islands only upon becoming sentence modifiers. The difficulty here is that it is hard to see how they could be analyzed as complex NP's.

(1.35) a. I told John that Bill was sick.
b. Bill was sick, I told John.
c. It's me who supposed that John is sick.
d. *It's me who John is sick, supposes.

This claim is independently supported by the unambiguity of (1.36c), which cannot be read as a paraphrase of (1.36b). Nevertheless, even though the parenthetical seems to be a deep, rather than a surface, structure modifier here (see Nobel, 1971), there is still no natural way of analyzing it as a complex NP.

(1.36) a. I suppose John is sick.
b. John is sick, I suppose.
c. It's me who supposes that John is sick.
d. *It's me who John is sick, supposes.

A second case which raises problems for the CNPC-reductionist hypothesis concerns the comparative and equitative constructions. The source of such constructions is one of the least well understood problems in transformational grammar,[3] but I shall assume an analysis

---

[3] Ross (1969a) proposed (ib) as the source of (ia), in an attempt to explain the occurrence of any or ever inside comparative clauses, as in (iia), as well as the ill-formedness of comparative clauses containing overt negation, as shown by (iib).

(i) a. John is taller than Bill.
b. John is tall to an extent to which Bill is not tall.

(ii) a. John is richer than any of my friends has ever been.
b. *John is richer than Bill isn't.

In his dissertation, Ross no longer defended (ib) as the source of (ia), as he had noticed that equitatives exhibit the same properties as comparatives, as can be seen by comparing (ii) and (iii).

(iii) a. John is as rich as any of my friends has ever been.
b. *John is as rich as Bill isn't.

As there is no obvious way in which equitatives can be analyzed as containing negation, Ross presumably concluded that the facts of (iia) and (iib) are not determined by negation and therefore do not force an analysis of (ia) along the lines of (ib).

There is also a second difficulty with Ross' analysis in (ib); (ia) and (ib) are not paraphrases. Thus, if John is taller than Bill, Bill is necessarily not taller than John, while if John is tall to an extent to which Bill is not, Bill necessarily is tall to an extent to which John is not; in other words the relation between John and Bill is asymmetric in (ia) and symmetric in (ib).

I believe, however, that Ross' initial analysis of comparatives made an important point, namely, that comparatives should be analyzed as containing negation at some level of representation, although not necessarily as in (ib).

One fact which suggests that comparatives are inherently negative is the manner in which they interact with Coordination-Reduction.

Consider first (iv) and (v).

(iv) a. John thinks that Mary is pretty and John thinks that Jill is ugly.
b. John thinks that Mary is pretty and that Jill is ugly.
c. John thinks that Mary is pretty or that Jill is ugly.

(v) a. John doesn't think that Mary is pretty and John doesn't think that Jill is ugly.
b. John doesn't think that Mary is pretty and that Jill is ugly.
c. John doesn't think that Mary is pretty or that Jill is ugly.

The reduced paraphrase of (iva) is (ivb), not (ivc), while the reduced paraphrase of (va) is (vc), rather than (vb). The principle which determines the change from and to or in (v) is known as DeMorgan's laws. I shall now show that DeMorgan's laws operate in comparatives, but not in equitatives, which strongly suggests that the former, but not the latter, are inherently negative. Thus, consider (vi) and (vii).

(vi) a. John is as clever as Mary is astute and John is as clever as Jill is sly.
b. John is as clever as Mary is astute and as Jill is sly.

(vi) c. John is as clever as Mary is astute or as Jill is sly.

(vii) a. John is cleverer than Mary is astute and John is cleverer than Jill is sly.
b. *John is cleverer than Mary is astute and than Jill is sly.
c. John is cleverer than Mary is astute or than Jill is sly.

In (vi), as in (iv), the reduced form of the a-sentence is the b-one, while in (vii), as in (v) the reduced form of the a-sentence is the c-one. It is difficult to see what, outside of DeMorgan's laws, could determine the difference between the paradigm in (vi) and the one in (vii).

Another piece of evidence, weaker than the first one, but which nevertheless suggests that comparatives are inherently negative, is provided by the surface structure of French comparatives and equitatives. In French, if the verb of the comparative or equitative clause has not been deleted, there must be a negative morpheme in the former case, but there cannot be one in the latter, as shown in (viiib) and (ixb), the French counterparts of (viiia) and (xia) respectively.

(viii) a. John is taller than George is.
b. Jean est plus grand que Georges $\left\{\begin{matrix} \text{ne} \\ {*}\_\_ \end{matrix}\right\}$ l'est.

(ix) a. John is as tall as George is.
b. Jean est aussi grand que Georges $\left\{\begin{matrix} {*}\text{ne} \\ \_\_ \end{matrix}\right\}$ l'est.

The acceptable version of (viiib) is a paraphrase of (viiia), not of the semantically ill-formed (see below) *_John is taller than George isn't_. The particle _ne_ has no semantic import, which is shown by the paraphrase relation between (viiia) and (viiib), as well as by the fact that semantic _ne_ always has a reduplicated clause-mate, usually _pas_, while no such reduplication is possible in (viiib), as shown by the ungrammaticality of *_Jean est plus grand que Georges no l'est pas_. Rather, _ne_ is probably itself a reduplication of some semantically negative morpheme in the string, which seems quite plausible in view of the fact that French has a rule reduplicating negation in any case (although the reduplicative rule in (viiib) is probably not the same as the rule which yields _Jean n'est pas ici_, since the reduplicated form is necessarily a clause-mate of the original in the latter case, but not in the former). If _ne_ in (viiib) is a reduplicated form, it is hard to see what the semantically negative element in the string is, outside of _plus_ 'more'.

The import of the above discussion is that negation should not appear inside the comparative clause, as in (ib), but rather in the semantic representation of _more_. I have not found, however, a satisfactory way of doing this. One possibility would be to represent _more (than)_ as _NOT EQUAL (TO)_. However, this representation would be incomplete, for *not equal to* is a symmetric relation, while *more than* is asymmetric; in fact, *not equal to* is equivalent to *more than* or *less than*, so that a semantic feature like [-EQUAL] could appear in the semantic representation of both *more* and *less*, provided that there is another feature which distinguishes between the two relations, by having the value + in one and - in the other; the problem is that it is not obvious what that feature should be. It should be noted, incidentally, that *less* has the same inherently negative semantic properties as *more*; this can be verified by substituting *less* for *more* and *moins* for *plus* in the above examples.

Various generative semanticists have proposed that (ia) be represented as (x).

(x) John is tall to an extent which exceeds the extent to which Bill is tall.

This analysis has certain difficulties which I shall touch upon further down in the text, but one rather interesting difficulty which, as far as I know, has not been noticed so far, is that although *exceed* is synonymous with *be more than*, the latter only is inherently negative. Thus, *exceed* exhibits the paradigm in (xi), which, surprisingly enough, is similar to the paradigm in (vi) rather than to the one in (vii), since the b-, rather than the c-sentence, is a paraphrase of the a-one.

(xi) a. John exceeds Mary in intelligence and John exceeds Bill in temerity.
b. John exceeds Mary in intelligence and Bill in temerity.
c. John exceeds Mary in intelligence or Bill in temerity.

In conclusion, I have been unable to discover an adequate semantic representation for *more than* and *less than*, beyond the observation that they must somehow involve negation.

It is also worthwhile to evaluate the evidence in (ii), which led Ross to positing (ib) as the source of (ia). (iia), is, in all probability, irrelevant to negation, in view of the acceptability of (iiia); the conclusion seems rather unavoidable that *any* and *ever* occur in more than just negative environments. Concerning (iib), Ross implied that it was out for the same reason as **John didn't see nothing* is out in standard English; in other words, he attributed the badness of (iib) to a surface property of English. However, I believe that (iib), as well as (iiib), is out on semantic grounds. Thus, only scalar properties can be compared, and (at least certain) negated adjectives do not stand for scalar properties; this can be seen by considering the ill-formedness of **to a very high degree, John isn't clever*, **John is exceedingsly not clever*, etc. In short, I am claiming that (iib) and (iiib) are bad for the same reason for which **John is cleverer than Bill is an assistant professor* is bad.

consonant with Geis' reductionist hypothesis. Thus, I shall assume that the source of (1.37a) is roughly (1.37b), and that the source of (1.37c) is roughly (1.37d); these assumptions make it possible to exclude (1.38a) and (1.38b) by the CNPC.

(1.37) a. John loves Mary more than he hates Bill.
b. John loves Mary to a degree which exceeds the degree to which he hates Bill.
c. John loves Mary as much as he hates Bill.
d. John loves Mary to a degree which {is the same as / equals} the degree to which he hates Bill.

(1.38) a. *It's Bill who John loves Mary more than he hates.
b. *It's Bill who John loves Mary as much as he hates.

In Chapter Six of his thesis, Ross points out that feature-changing rules are also subject to the CNPC. However, although the rule which permits the occurrence of *any* or *ever* can go into a comparative clause, it cannot go into what the reductionist hypothesis suggests as its source, as shown in (1.39).

(1.39) a. John is richer than any of my friends has ever been.
b. *John is rich to an extent which exceeds the extent to which any of my friends has ever been rich.

In order to explain both (1.38) and (1.39), it becomes necessary to assume that the transformation which replaces *an extent which exceeds the extent to which* by *more than* *precedes* feature-changing rules, but *follows* chopping rules. Unfortunately, there is evidence of another sort which requires that the transformation in question *precede* chopping rules. Indeed, elements of comparative clauses can be chopped following Comparative Deletion, as shown by (1.40).

(1.40) a. John is taller than a man.
b. I know a man who John is taller than.
c. *John is tall to an extent which exceeds the extent to which a man.

If the lexical transformation follows the chopping rules, (1.40c) would become the input to Relativization, and the latter would be blocked by the CNPC, since *a man* is inside a relative clause; therefore, we would wrongly predict that (1.40b) is unacceptable. Thus (1.38) requires that the lexical transformation follow Relativization, while (1.40) requires the opposite ordering; the resulting paradox suggests that the lexical transformation we have posited does not exist, and therefore that (1.37b) and (1.37d) are not the sources of (1.37a) and (1.37c) respectively.

A third case which creates rather serious problems for the reductionist hypothesis is illustrated by the contrast in grammaticality between the members of the pair in (1.41).

(1.41) a. (?)The accident which a reporter has just disclosed the place of took five lives.
b. *The accident which the police found Mary's body at the place of took five lives.

In both (1.41a) and (1.41b) an element of the construction the place of the accident has been relativized. Regardless of whether that construction can be analyzed as a complex NP or not, there is no reason to believe that the transformation(s) by virtue of which it assumes its surface form is ordered before Relativization in (1.41a) but after Relativization in (1.41b). Thus, unless we refer to the fact that the place of the accident is an adverbial in (1.41b) but not in (1.41a), there seems to be no way of accounting for the difference in acceptability between these two sentences.

A similar difficulty is created by reduced comparative clauses. As (1.40b) shows, the adverbial (without its connector) can be moved, but elements of that adverbial cannot, as shown by the contrast in acceptability between (1.42b) and (1.42c).

(1.42) a. John is taller than a member of the committee.
b. This is the member of the committee who John is taller than.
c. *This is the committee which John is taller than a member of.

Paradigms similar to (1.42) can be constructed using other reduced adverbial clauses, as shown in (1.43) and (1.44).

(1.43) a. John left the city because of a brother of the queen.
b. ?This is the brother of the queen who John left the city because of.
c. *This is the queen who John left the city because of a brother of.

(1.44) a. John left the city in spite of a brother of the queen.
b. ?This is the brother of the queen who John left the city in spite of.
c. *This is the queen who John left the city in spite of a brother of.

To summarize: the three types of islands[4] we have examined above, external, and an element internal, to it.

---

[4] This term was coined by Ross (1967, Ch. 6); an island is a configuration such that certain rules cannot involve an element

---

namely, parentheticals, comparative and equitative clauses, and nonclausal adverbials, have been shown not to be reducible to the CNPC-case, and thus suggest rather strongly that, given Ross' framework, observational adequacy requires the imposition of a fifth constraint over and above the CNPC, the CSC, the LBC, and the SSC, which we may call the Optional Adverbial Constraint (or the OAC), and whose statement would have to be something like (1.45).

(1.45) The OAC: No element of an optional adverbial[5] (i.e., a VP- or S-modifier) can be moved from under the dominance of the S-node which immediately dominates the node ADV (or the corresponding node NP, in a grammar which does not employ the symbol ADV).

---

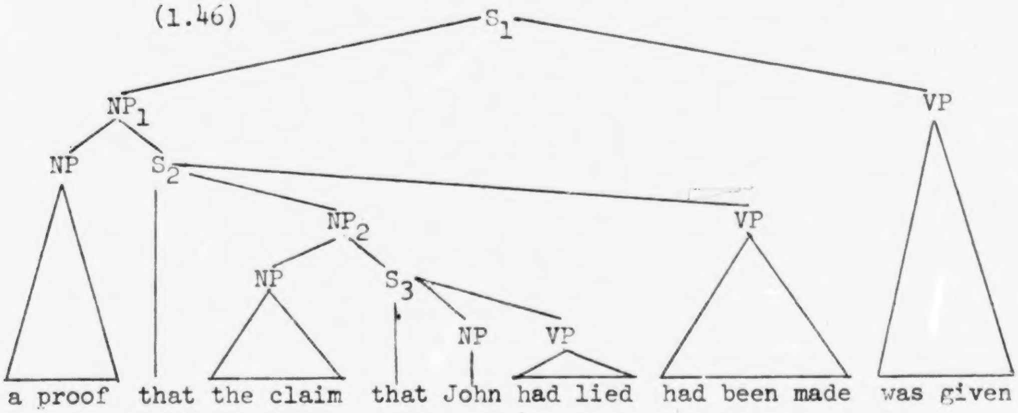
[5] "Adverbial" is here used in the narrower sense, i.e., not including the connector.

It is rather easy to see that the three cases which required (1.45) cannot be disposed of by the A-over-A Principle either. Indeed, the migration of NP's from parentheticals could not be blocked, for there is no reason to posit a node NP over the parentheticals; the blocking of movements out of comparatives and equitatives would, of course, depend on whether such clauses are dominated by the category NP or not; as for cases like (1.41), the contrast in acceptability between the a and the b sentence must go unexplained.

In the alternative analysis I will develop in Chapters Three and Four, I shall argue that the OAC is an unnecessary addition to the grammar, and that the three cases which seemed to argue for its introduction follow automatically from more general principles of analysis.

1.2.4.2. In addition to the four constraints Ross proposes as substitutes for the A-over-A Principle, he also puts forward two constraints limiting the scope of movement transformations in general. To see why that was necessary, we shall consider how case (1.15c) of the A-over-A principle is handled in Ross' framework. In section 2.3, Ross proposes to account for the unacceptability of (1.18) with the CNPC. Indeed, assuming that the stage in the derivation of (1.18) which precedes Extraposition-from-NP is (1.46), we can see that in order to derive (1.18), we must move S_3--an element of S_2--out of NP_1, and the latter is a Complex NP whose modifying clause is S_2.

(1.46)



However, if NP_1 were not a Complex NP, S_3 would still be prevented from extraposing "too far", as (1.47b) cannot be derived from (1.47a).

(1.47) a. That the claim that John had lied had been made was obvious.
b. *That the claim had been made was obvious that John had lied.

It seems rather counterintuitive to claim that Extraposition and Extraposition-from-NP are constrained by different principles, and Ross proposes, in Chapter Five, that the notion of *bounding* be added to the theory of grammar. Thus, a rule is "upward bounded" if its domain is restricted to whatever is dominated by the "first sentence up". Ross further proposes a putatively universal constraint, the *Right Roof Constraint* (henceforth, the *RRC*), which easily accounts for (1.18) and (1.47b).

(1.48) The *RRC*: Rules which move elements to the right are upward bounded.

The RRC severely limits rightward movements. While leftward movements are not, in general, limited in scope, they are nevertheless limited in "direction", for no element can move out of its "strip" (or "highest island"). The highest island of an element Y is the set of all elements which command Y (the notion *command* was developed by Langacker, and, informally, asserts that X commands Y if X is in a sentence which dominates Y). Thus, it turns out that no movement transformation is completely free, and Ross proposes an additional restriction which he called elsewhere the *Highest Island Constraint* (henceforth, the *HIC*), and which can be formulated as follows:

(1.49) The *HIC*: A transformation can only move an element to a position which commands that element's initial position.

1.2.5. In Chapter Six of his dissertation, Ross points out that chopping rules are not the only ones which are subject to his various constraints and that feature-changing rules and some deletion rules also have this property. For example, (1.50a), is claimed to be bad because the rule which changes some to any reaches into a relative clause from outside the containing complex NP, and (1.50b) is claimed to be bad because a deletion rule has committed the same offense.

(1.50) a. *I never told you I had seen the man who ever fought against anybody.
b. *Jill is prettier than you are aware of the fact that Mary is.

On the other hand, Ross notices that not all unbounded rules are subject to island constraints. For example, copying rules, pronominalization rules, and certain deletion rules are not, as shown by the acceptability of (1.51a)-(1.51c), in which the respective rules reach inside complex NP's in the same way in which feature-changing and deletion rules do in (1.50).

(1.51) a. Who did I show you a girl who likes him? [dialectally restricted]
b. John dislikes all the girls who have chased him.
c. You can certainly use a rifle, but I know a girl who thinks you can't.

Ross attempts to make sense of the situation by means of the following statement, which I shall call the Dichotomous Behavior Principle (henceforth: the DBP):

(1.52) The DBP: Chopping rules, feature-changing rules, and unidirectional rules of pronominalization obey the CNPC, CSC, LBC and SSC; bidirectional rules of pronominalization and copying rules do not.

The following two claims made by the DBP (the first, implicitly, the second, explicitly) are worthy of attention:

(1.53) a. Deletion is a subcase of pronominalization, and both phenomena behave alike;
b. The criterion which determines the applicability of island phenomena to specific rules of pronominalization is whether the rules at issue operate unidirectionally or bidirectionally.

Thus, the difference in acceptability between (1.54a) and (1.54b) must be accounted for by arguing that the rule which introduces the underscored pronoun in (1.54a) is the rule of pronominalization which operates in other environments, while the rule which deletes that pronoun to yield (1.54b) is a special deletion rule restricted to this specific environment.

(1.54) a. The mink coat is ready for you to find a girl willing to put it on.
b. *The mink coat is ready for you to find a girl willing to put on.

Notice that the only reason for saying that the deletion rule is unidirectional is that the for-to clause cannot prepose; the non-existence of a preposing rule does not lead Ross to the claim that the pronominalization rule which yields (1.54a) is unidirectional.

Observe now that the rule which yields (1.50b) is clearly the same rule which yields (1.55a) and (1.55b), that is, VP-Deletion.

(1.55) a. Although her brother is, I am afraid that Mary is not efficient enough.
b. Jill is conceited, but I don't really know a girl who isn't.

To make this proposal work, Ross would be forced to claim that the rule of pronominalization which yields (1.54a) is the same rule which pronominalizes in other environments, while the VP-Deletion rule which operates in comparative constructions like (1.50b) is different from the VP-Deletion rule which operates in non-comparative environments, since the latter is (i) bidirectional, and (ii) exempt from island-constraints, as shown by (1.55a) and (1.55b), respectively. Clearly, such a position is inconsistent.

In connection with the DBP, Ross writes, "It is at present a total mystery ... why unidirectional pronominalizations should obey the constraints..." However, no mystery is involved, since the distinction Ross makes between unidirectional and bidirectional pronominalization is almost certainly irrelevant. Thus, if we examine (1.54) more carefully, it becomes apparent that its unacceptability not only has nothing to do with unidirectionality, but has nothing to do with deletion in the first place. The relevance of deletion to comparative formation[6] was probably assumed because identity between the 'comparator' and the 'comparee' requires the deletion of the latter in English (although not, for example, in Japanese). But it is not necessary to resort to Japanese in order to expose the irrelevance of deletion, for English has comparative constructions like (1.56), in which no deletion occurs.

---

[6] Chomsky (1971, note 29) also assumes the relevance of deletion in this case, and proposes that the comparee is in fact moved by cyclic hopping until it is adjacent to than, after which it is deleted. Thus, the island constraints found with comparative constructions are blamed on an invisible chopping rule. This proposal is untenable for the same reasons as Ross'.

(1.56) John is taller than it has been assumed that Bill is fat.

Nevertheless, constructions like (1.56) are subject to the CNPC, CSC, SSC and OAC, as (1.57a)-(1.57d) respectively show.

(1.57) a. *John is taller than I $\left\{\begin{array}{l}\text{know a boy who is fat}\\\text{am aware of the fact}\end{array}\right.$ that Bill is fat $\}$.

b. *John is taller than Bill is fat and employed full-time.

c. *John is taller than that Bill is fat is clear.

d. *John is taller than Mary left Bill because he is fat.

The problems raised by (1.57) for the DBP and for the Transformational Position in general will be taken up in the next section. Suffice to point out at this stage that they strongly suggest that VP-Deletion in (1.50b) is not the factor responsible for unacceptability.

There remain, however, a number of puzzling cases; thus, VP-Deletion, S-Deletion, Sluicing, Super-Equi-NP-Deletion, and perhaps other rules, are exempt from the constraints, while Japanese Relativization, the rule which transforms (1.54a) into (1.54b), and the rule which transforms (1.58a) into (1.58b) are not.

(1.58) a. This rock is too heavy for you to find a boy willing to pick it up.

b. *This rock is too heavy for you to find a boy willing to pick up.

I suggest that what distinguishes the two classes of deletion rules is that in one class of cases deletion applies to structures subject to an obligatory deep structure condition to the effect that the clause which contains the deletee must contain a coreferent of the deletor, while in the other class of cases no such condition exists, and consequently no deletion need occur in the corresponding structures. The pertinent constraint, which I will defend and propose an explanation for in Chapter Four, is (1.59).

(1.59) Deletion rules are subject to island constraints if and only if there exists an independent condition requiring that an element coreferential with the deletor occur within some configuration which contains the deletee.

(1.59) correctly predicts that Japanese Relativization is constrained by islands, because a relative clause is universally ill-formed if it does not contain a coreferent of the head at some level of representation; in contrast, rules like VP-Deletion are not coupled with such conditions, as can be seen by considering

(1.60), which is identical with (1.55b) in all respects except for the fact that the identical VP's in the latter are not identical in the former, and no deletion consequently takes place.

(1.60) Jill is conceited, but I don't really know a girl who isn't in love with herself.

With respect to (1.54) and (1.58), it will be shown in Chapter Four that the occurrence of a coreferent of the subject of the main verb within the for-to clause is an inference 'invited' by a semantic condition requiring that a resultative relation obtain between the main and the for-to clauses; this invited inference (for a definition of this term, see Geis and Zwicky (1971)) has in general the same effects as a strictly linguistic condition, an interesting but not isolated fact (I will show below other instances of syntactic form being determined by invited inferences or by 'conversational implicatures' in the sense of Grice (1968)).

It should be noted that (1.59) refers to strict deletion, not to pronominalization in general, and therefore predicts that only the b-, but not the a-sentences in (1.54) and (1.58) are bad.

1.2.6. It is important to notice at this point that if (1.59) is correct, the Transformational Position is observationally inadequate.[7] The reason is that the conditions which dictate or invite the inference of the occurrence of two coreferential NP's are semantic, and consequently belong in deep structure. On the other hand, the pertinent deletion processes are probably very late ones, and belong to a stage very close to surface structure. Therefore, an adequate statement of (1.59) within the theoretical framework which Ross espouses (Generative Semantics) requires the employment of a global rule.

The claim that island constraints should be reformulated as global rules has in fact been made by Ross himself in Ross (1969), where it was pointed out that there is a deletion rule, Sluicing, which, when deleting the remainder of an island out of which an element had been moved by Question-Movement, yielded more acceptable sentences than those which would have resulted had Sluicing not applied. The point is demonstrated in (1.61) for the CNPC, CSC, SSC and OAC.

(1.61) a. ?I know you kissed a girl who had received a coat from someone, but I don't know who (*you kissed a girl who had received a coat from).
b. ?Bill and someone were dancing, but I don't know who (*Bill and were dancing).

---

[7] I am assuming a transformational grammar which does not allow arbitrary coding symbols. If such symbols are tolerated, global rules can of course be avoided.

(1.61) c. ?That John loves someone is well-known, but I don't know who (*that John loves is well-known).

d. ?John left Mary because she had slept with someone, but I don't know who (*John left Mary because she had slept with).

Because of facts like those in (1.61), Ross proposed the following ammended formulation of island constraints:

(1.62) When an element is moved out of an island, unacceptability results; if the remainder of the island is later deleted, unacceptability of lesser severity results.

Clearly, both the reduction in acceptability due to island-deletion and unacceptability due to obligatory coreferentiality coupled with deletion must be global conditions on island constraints.

With regard to (1.59), one would want to state it more generally than just for deletion rules, since there is a definite parallelism between the fact that non-null pronominalization does not create unacceptability even in the presence of an obligatory coreferentiality condition, while strict deletion does, and the fact that copying rules do not create unacceptability, while chopping ones do. In both cases, no unacceptability arises when the affected element is reduced to a non-null pro-form, and unacceptability does arise when the affected element disappears completely. The similarity between the two sets of phenomena can be captured more perspicuously if we regard chopping rules as proceeding in two stages, i.e., copying followed by deletion (as proposed, for example, in Drachman (1970)). Such a view is particularly reasonable for rules like English Relativization or Question-Movement, which do not appear as copying ones in the standard dialect, but which nevertheless are realized as copying rules in many non-standard dialects or even in relaxed styles of standard-dialect speakers; it is to be assumed that children exposed to both possibilities will formulate the rules in question as copying ones, followed by deletion in specific situations. If so, we may notice that copying rules resemble the deep structure coreferentiality condition in that they also create two necessary coreferents. The similarity between chopping and deletion under obligatory identity can now be made explicit: in both cases, strict deletion occurs, and in both cases an element coreferential with the one acting as deletor must exist.

The possibility of capturing the generalization just stated with global rules depends on the constraints placed on global rules. If we require that global rules refer to specific stages in a derivation, the generalization cannot be captured, for in one instance it is necessary to refer to deep structure and in the other to the output of copying. If we do not require global rules to be specific in that sense, the generalization can be captured.

It is noteworthy that there are a great deal more global conditions which need to be imposed on island constraints. One such condition is necessary to account for the ill-formedness of the sentences in (1.57), which clearly cannot be blamed on some movement or deletion transformation. One may be tempted to impose a deep structure condition to the effect that the comparee cannot occur in an island if the comparator is outside that island (equitative constructions work in the same way, as the reader can convince himself by replacing taller than with as tall as in (1.56) and (1.57)). However, the pertinent condition cannot be a deep structure one,[8] because (1.63), which has the same deep structure as (1.57c), is acceptable.

(1.63) John is taller than it is clear that Bill is fat.

Also, the well-formedness of comparatives seems to depend on certain conditions of formal similarity between the compared elements (I have not investigated these conditions in detail, as their precise statement undoubtedly warrants a separate study; suffice to notice, however, that such conditions exist). Thus, consider the difference in acceptability between (1.64a) and (1.65a), as well as between (1.64b) and (1.65b), which have been claimed to be related transformationally.

(1.64) a. It seems that Bill is clever as much as it seems that Joe is stupid.
b. *John is as willing to please as for anyone to offend Mary is easy.

(1.65) a. ?*It seems that Bill is clever as much as Joe seems to be stupid.
b. John is as willing to please as Mary is easy to offend.

(1.65a) shows that an acceptable deep structure becomes bad by a transformation (Subject-Raising), and (1.65b) shows that a bad deep structure becomes acceptable by a transformation (Tough-Movement). It would seem that the well-formedness of comparatives and equitatives must be determined by a shallow structure condition (in the sense of some level which precedes post-cycle deletions) requiring some kind of formal similarity between the compared VP's (the condition cannot be a surface one, since formal similarity cannot be determined from the surface in sentences like John is as tall as you think Bill is).

The shallow structure condition suggested above is not sufficient to determine the acceptability of comparatives or equitatives as the latter also depends on the semantic compatibility of the compared phrases. Thus, (1.66a) and (1.66b) are all right, but (1.66c) and (1.66d) are not.

---

[8] In Grosu (1972a), it was mistakenly claimed that the pertinent condition is a deep structural one.

(1.66) a. John took to his heels as swiftly as Bill fought bravely.
b. Mary is as pretty as Jill is refined.
c. *John took to his heels as swiftly as Jill is refined.
d. *Mary is as pretty as Bill fought bravely.

As semantic compatibility belongs in deep structure, it appears that the determination of the well-formedness of comparatives and equitatives requires at least the following statement:

(1.67) Comparatives and equitatives are ill-formed if (i) the compared elements are semantically incompatible, or (ii) the comparator is external and the comparee is internal to some island configuration in shallow structure.

Notice that the two conditions of (1.67) are disjoined, and there is therefore no motivation for regarding (1.67) as a global rule, but rather as a set of two independent conditions. Consequently, the island constraints must bear a third condition, but a shallow-structure rather than global one, namely (1.67ii).

It should be pointed out that there is an alternative way of accounting for the ungrammaticality of the sentences in (1.57) without recourse to (1.67). Thus, given an analysis of comparative and equitative constructions along the lines suggested by M. Geis (and other generative semanticists), as in (1.37), one can blame the badness of the sentences in (1.57) on a movement transformation. For example, if an earlier stage in the derivation of (1.57b) is **John is tall to an extent which exceeds the extent* [*Bill is* [*fat to that extent and an assistant professor*]], Relativization will move the rightmost doubly underscored one in violation of the CSC.

The above account is certainly a possible one; however, I have not adopted it because it requires strong ordering assumptions, which do not have enough independent support. Specifically, it requires that Relativization apply to the degree adverb before the lexical insertion rule which replaces *to an extent which exceeds the extent to which* by *more than*; this lexical insertion rule must itself precede the rule of Comparative Deletion which turns *John is brighter than Bill is* into John is brighter than Bill, for this rule can apply only when the comparee is immediately embedded to the comparator, as shown by the ungrammaticality of **John is brighter than I told you that Bill*; finally, this rule must in turn precede Relativization, in order to allow *I know a man who John is taller than* and to block **I know a man who John is taller than is*. All this could be done by making Relativization, Comparative Deletion, and the lexical insertion rule cyclic, an ordering for which I do not know of much independent justification (in particular, lexical insertion has usually been claimed to take place at the end of the cycle).

It should also be pointed out that if Relativization and lexical insertion are made cyclic, Geis' proposal that the island character of adverbials be explained by the CNPC collapses, for no adverbial would be a complex NP at the stage at which Relativization moves an element out of it. In other words, Geis could not simultaneously maintain his claims that the restrictions on adverbials are reducible to the CNPC and that (1.37a) and (1.37c) are derived from (1.37b) and (1.37d), respectively (notice that Relativization and lexical insertion would have to be cyclic not only in order to account for constructions containing comparatives, such as I know a man who John is taller than, but also in order to account for constructions containing other adverbials, such as this is the woman who John committee suicide because of, and this is the girl who John became a priest in spite of).

The various problems mentioned above had led me to adopt the view that the sentences in (1.57) are bad because the semantic comparative relation goes into an island, rather than because of the movement of a degree adverb out of that island (even though the latter analysis is not untenable, provided that we discard the claim that adverbs are islands by the CNPC). It should be noted that ascription of island-sensitivity to semantic relations is needed independently of comparatives or equitatives; D. James (1972) has discovered quite recently that the relation between interjections and their scope is also subject to island constraints, and it is highly improbable that an analysis involving movement could be devised for the cases she considered.

Another condition to be imposed on island constraints concerns the distribution of neutralized elements like any or ever. As Ross pointed out, such elements, which must be commanded by some neutralizer (e.g., negation), cannot occur inside islands if the neutralizer is outside the island. This restriction is shown in relation to the CNPC, CSC, SSC and OAC in (1.68)-(1.71) respectively.

(1.68) *I never said that you met a boy who ever harmed anyone.

(1.69) *I never said that Bill is brave and Mary ever harmed anyone.

(1.70) *I never said that for you to ever harm anyone is desirable.

(1.71) *I never said that Bill loves Mary, because she ever harmed anyone.

At the time Ross wrote his thesis, it was assumed that any is introduced by a rule of Indefinite-Incorporation (or Some-Any Suppletion), and Ross argued that that rule is subject to island constraints. This assumption was later questioned by Robin Lakoff (1969), but independently of that fact, it is hard to see what rule could be blamed for the occurrence of ever inside islands. Consequently, the occurrence of ever, and probably any, must be determined by some condition on a level. There is no reason that I can see why this condition could not be stated at the same level as (1.67ii), that is, shallow structure, where the distribution of neutralized elements in general could be stated. There is, however, the problem that the level of shallow structure has not yet been defined precisely enough. Thus, certain writers have suggested that the output of the cycle be regarded as shallow structure, but such a definition wouldn't work if Extraposition is a cyclic rule, for the well-formedness of sentences containing neutralized elements must be determined after sentential subjects have extraposed. Moreover, if Question-Formation is post-cyclic, it must also precede shallow structure, as shown by (1.72)

(1.72) I never heard about Bill's murdering $\left\{\begin{array}{l}\text{Mary}\\ \text{anyone}\end{array}\right\}$.

(1.73) Whose murdering $\left\{\begin{array}{l}\text{Mary}\\ \text{*anyone}\end{array}\right\}$ did I never mention to you?

Shallow structure cannot, however, follow all movement rules. In particular, it cannot follow Topicalization (as the acceptability of (1.74) shows), and it cannot follow Pseudo-Clefting.

(1.74) That Bill ever harmed anyone his sister just can't believe.

With respect to Pseudo-Clefting, there have been proposed at least three competing analyses, and I shall argue at great length in 2.1.3 that only an analysis which involves the "extraction" of the pseudo-cleft element from a relative clause in subject position is tenable. Suffice to point out at this stage that the cleft constituent must still be within the complex-NP subject in shallow structure if the ungrammaticality of (1.75a) and (1.75b) is to be accounted for by (1.67ii).

(1.75) a. *Bill is taller than what I told you is that Bill is fat.
b. *I never told you that what you fear most is that Mary ever harmed anyone.

Notice that the unacceptability of (1.75a)-(1.75b) has to be accounted for by some condition like the CNPC, for predicates in construction with the copula are not islands in general, as shown by (1.76a)-(1.76b).

(1.76) a. This knife is sharper than what you saw yesterday is heavy.
b. I never claimed that what you told me last year was ever surprising to anyone.

I will not pursue the discussion of shallow structure any further, and will content myself with having pointed out that in all probability some post-cyclic rules must precede, while others must follow, shallow structure.

It is important, however, to ask whether Pseudo-Clefting can also follow all the unbounded movement rules, for those cannot touch pseudo-cleft constituents, as (1.77)-(1.79) show.

(1.77) *Who do you think that what John wants is to marry?

(1.78) *The girl who I think what John wants is to marry is sick.

(1.79) *Jill, I think what John wants is to marry.

The unacceptability of (1.77)-(1.79) would be explained if Question-Formation, Relativization and Topicalization preceded Pseudo-Clefting, for the pseudo-cleft would still be inside an island at that point. An account along these lines has in fact been attempted by Ross in connection with the island properties of clauses moved by Extraposition-from-NP. Thus, Ross proposed that the badness of sentences like those in (1.80)-(1.82) be explained by ordering Extraposition-from-NP after all unbounded movement rules.

(1.80) *I don't think that the man has been hired who ever robbed anyone.

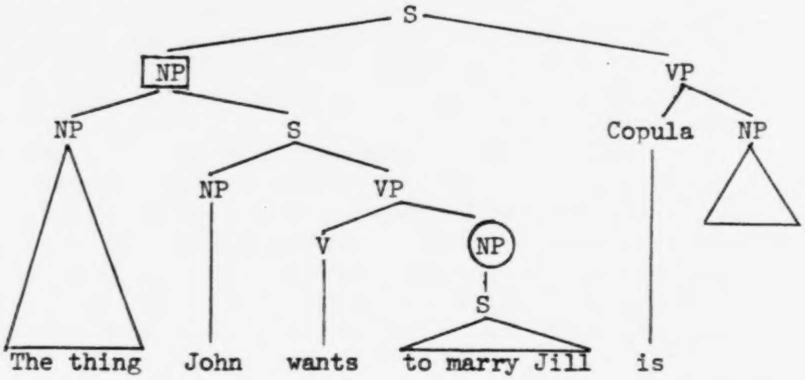
(1.81) *John is taller than a boy just left who was tall.

(1.82) *Which hat has a boy just left who was wearing?

Unfortunately, a similar solution is impossible for the sentences in (1.77)-(1.79). If (1.84) is the input to Pseudo-clefting in the derivation of (1.83) (for discussion, see section 2.1.3), it is clear that the extraction of to marry Jill must precede Relativization, for it is only after extraction that the copy left behind, namely something, becomes identical with the head of the relative clause.

(1.83) What John wants is to marry Jill.

(1.84)



Now, in order to explain the badness of (1.85) by the CNPC within the Transformational Position, we would have to claim that the lower occurrence of Jill is relativized while the circled NP is still dominated by the boxed NP.

(1.85) *Jill, who what John wants is to marry, is my cousin.

But this is impossible whether Relativization is cyclic or post-cyclic. If Relativization is cyclic (by which I mean that it is triggered as soon as the head of the relative clause is reached, and not that the relative pronoun 'hops' one cycle at a time, as suggested in Jackendoff (1969)), then Jill must be relativized before the circled NP in (1.84); in other words, Relativization on a higher cycle would have to precede Relativization on a lower cycle, with Pseudo-Clefting ordered in between. This would mean a genuine anti-cyclic ordering of Relativization (not merely its sequential application from the top to the bottom of the tree), and since no one has ever offered any evidence in support of such a position, we need not take it seriously. If Relativization is post-cyclic, which means that all instances of Relativization must apply in block (i.e., with no rules ordered between various applications thereof), and if Pseudo-Clefting precedes one instance of Relativization, it must precede all instances of Relativization; consequently, Jill cannot be relativized before the circled NP is pseudo-clefted.

The only way in which (1.78) and (1.85) can be excluded by the CNPC is through a fourth global condition, which I state tentatively as (1.86).

(1.86) Elements which cannot be moved because they belong in an island, cannot be moved even after they are no longer part of that island.

(1.86) can accomplish the desired results with respect to (1.77)-(1.79) and (1.85), but it cannot capture the intuition that (1.80)-(1.82) is bad for the same reasons as (1.77)-(1.79) or (1.85), namely that an element is moved out of a constituent which belonged to an island in remote representation. In other words, the handling of (1.80)-(1.82) by extrinsic rule ordering and of (1.77)-(1.79) and (1.85) by a global condition seems to say that the two situations are unrelated.

I will suggest in Chapter Four how the similarity between (1.77)-(1.79), (1.80)-(1.82) and (1.85) can be captured within the Behavioral Position, but for the moment we may adopt (1.86), which I think is observationally correct. We may notice that (1.86) has at least one interesting consequence, namely, that the level of shallow structure referred to in (1.67ii) cannot exist. To show this, we must consider the properties implicitly or explicitly ascribed to that level. The first one (1.87i) is definitional:

(1.87) (i) Shallow structure is not earlier than the output of the cycle.

The second property follows from (1.87ii) and our subsequent discussion:

(1.87) (ii) Shallow structure is the level at which the correct distribution of neutralized elements (i.e., comparees, _any_, _ever_, etc.) is stated.

The third property follows from (1.72)-(1.73):

(1.87)(iii) Question-Formation precedes shallow structure.

The fourth property follows from the well-formedness of (1.88) (in some dialects) and of (1.89) (in all dialects); with respect to the latter, its relevance depends on our accepting Akmajian's claim (1970) that clefts are derived from pseudo-clefts:

(1.88) What do you think that what John prefers is?

(1.89) What is it that you think John prefers?

(1.87) (iv) Question-Formation follows shallow structure.

Since in both (1.72)-(1.73) and (1.88)-(1.89) Question-Formation applies last- or post-cyclically, both applications take place at the same stage and no level can exist in between. I conclude that a level with the properties (1.87i)-(1.87iv) cannot exist.[9]

---

[9] Notice that we can use an argument of the same type for arguing that the rule which derives clefts from pseudo-clefts is not Extraposition, as Akmajian claimed. Indeed, (1.89) shows that Question-Formation must follow Akmajian's 'Extraposition'; but the difference in acceptability between (i) and (ii) shows that Question-Formation must follow Extraposition, or (ii) will be marked as ill formed by the SSC.

(i) *Who is that John loves clear?
(ii) Who is it clear that John loves?

As Akmajian's claim entails a cyclic ordering of Question-Formation and Extraposition within the last- or post-cycle, the rule which derives clefts from pseudo-clefts cannot be Extraposition.

The moral of this argument is not necessarily that a level of shallow structure cannot exist in principle, but that the distribution of neutralized elements cannot be stated at some well-defined level of linguistic structure.[10]

---

[10] Notice that even if the distribution of neutralized elements were stable in shallow structure, it would still be highly unnatural for island constraints to hold at that level for neutralized elements, but at different derivational stages in general. The reason is that island constraints can certainly not hold in shallow structure in general, for this level would then have to possess additional contradictory properties such as the following:

(1.87) (v) Pseudo-Clefting and Relativization follow shallow structure.
(vi) Pseudo-Clefting and Relativization precede shallow structure.

(1.87v) follows from (1.67ii) and the ungrammaticality of (1.71a) and (1.71b); (1.87vi) follows from the fact that Pseudo-Clefting and/or Relativization may not touch islands, as shown in (i).

(i) *What I told you about a boy who saw was the house.

The point I am trying to make is that it would be highly unnatural for a set of phenomena to constrain sometimes transformations and sometimes a given level.

We have so far noted four situations (in connection with (1.59), (1.62), (1.67), and (1.86)) which cannot be handled within the transformational Position. Two more such situations will be discussed in this section.

In his discussion of the CNPC, Ross (1967) pointed out the difference in acceptability between (1.90a) and (1.90b) (for most speakers, (1.90b) is intermediate in acceptability between (1.90a) and (1.90c), but Ross did not discuss this fact).

(1.90) a. *It's Mary I discussed the claim that Bill likes.
b. ?It's Mary I made the claim that Bill likes.
c. It's Mary I claimed that Bill likes.

Noticing the synonymy between (1.90b) and (1.90c), Ross suggested that the former be derived from the latter by a rule replacing claim with make the claim, which would follow all the post-cyclic unbounded movement rules, so that no island configuration should exist when the unbounded movement occurs. However, Ross pointed out that he could not provide a satisfactory solution within the Transformational Position, since the lexical replacement rule must also precede the various unbounded rules; the reason given was that those rules must be preceded by Passive, which must in turn be preceded by the lexical replacement rule, in order to allow the derivation of the *claim that John likes Mary has already been made by a number of people*. Of course, the ordering paradox noticed by Ross can be handled within the Derivational Position with a statement like the following:

(1.91) Violations of island constraints are considerably reduced in severity if the affected island was not an island in underlying representation.

The last problematic case for the Transformational Position I wish to discuss concerns the rule of (non-emphatic) Reflexivization. Ross claims that Reflexivization is subject to island constraints, because it cannot violate the CSC (the other constraints are inapplicable, because Reflexivization is both upward and downward bounded, i.e., the reflexive and its antecedent must be clause-mates). Ross supports his claim with the following examples:

(1.92) a. *Bill understands Mary and himself.
b. *Bill and Mary washed himself.
c. *Andy pinched Sarah and tickled herself.
d. *The gun and a description of itself lay on the table.

Notice that the above examples are relevant within the Transformational Position only if Coordination-Reduction precedes Reflexivization. If Coordination-Reduction follows Reflexivization, then the input to the former in the derivations of (1.92a)-(1.92d) will be (1.93a)-(1.93d).

(1.93) a. Bill understands Mary and Bill understands himself.
b. *Bill washed himself and Mary washed himself.
c. *Andy pinched Sarah and Andy tickled herself.
d. *The gun lay on the table and a description of itself lay on the table.

It can be seen that, given the ordering Reflexivization ← Coordination Reduction, the b-d sentences in (1.92) are bad independently of the CSC, and the badness of (1.92a) cannot be accounted for by the CSC, for the violated island does not yet exist at the stage at which Reflexivization takes place. Therefore, Ross' claim depends on the ordering Coordination Reduction - Reflexivization being the correct one.

Let us begin by contrasting the ill-formed (1.92a) with the well-formed (1.94).

(1.94) John admires Mary and despises himself.

In order to account for both (1.92a) and (1.94), Ross would be forced into the highly unnatural position that some instances of Coordination Reduction precede, while others follow, Reflexivization. This unnatural position could be avoided by regarding the two parts of the CSC (i.e., the one which refers to entire coordinate terms and the one which refers to parts of coordinate terms) as essentially independent restrictions (in Chapter Three I will argue that this is precisely the correct conclusion), and that Reflexivization is subject, along with a number of other pronominalization phenomena, only to the first part of the CSC. In that case, the acceptability of (1.94) need not trouble us.

It remains to decide on the ordering of Reflexivization and Coordination Reduction. Reflexivization must be cyclic, if the grammatical (1.95) is to be generated.

(1.95) Mary believes herself to have done harm to herself in her sleep.

Coordination Reduction must be either cyclic or post-cyclic but not precyclic) because it must follow the cyclic rule of Passive, as (1.96) suggests.

(1.96) John hit Mary and was hit by Bill.

Making the more conservative assumption that Coordination Reduction is cyclic, let us consider the derivation of (1.92a). Its source must be something like (1.97).

(1.97) Bill understands Mary and Bill understands Bill.

Since Reflexivization is cyclic, it will apply on the first cycle, yielding (1.93a); Coordination Reduction cannot apply on the first cycle, because its structural description is not met, and must therefore apply on the second cycle and to the output of Reflexivization. Consequently, the unacceptability of (1.92a) cannot be blamed on the CSC within the Transformational Position, and a sixth global condition like (1.98) seems to be called for.

(1.98) A rule subject to island constraints is blocked even when it affects a structure which is not internal to an island, if that structure becomes internal to an island after the application of the rule in question.

We have considered six non-transformational conditions on island constraints, namely (1.59), (1.62), (1.67), (1.86), (1.91) and (1.98), and have seen them to be necessary for observational adequacy purposes. There is no reason to believe that the list is closed, but even if it is, there are sufficient grounds for rejecting the Transformational Position as observationally inadequate.

With respect to the Derivational Position, we have seen that it was observationally adequate for the cases considered, and we may credit it with observational adequacy in handling other similar cases, if such should arise. On the other hand, there is no reason for believing that the Derivational Position can do more than give an observationally adequate account of the facts. The Derivational

Position could be claimed to capture an important generalization if the six global conditions I have listed could be shown to form a natural class. However, there is no way that I can see in which this can be done. (1.86) and (1.91) might suggest a generalization like 'a configuration counts as an island if and only if it has that status in underlying representation', but this biconditional statement is falsified both ways by (1.62) and (1.98); (1.62) shows that the violation of a configuration with island status in underlying representation becomes immaterial due to something which happens at a later stage, and (1.98) shows that a configuration with no island status in underlying representation can impose island constraints due to something that happens later on. Until someone can show that the list of global conditions given above forms a natural class, we can regard the Derivational Position as descriptively and explanatorily inadequate, since the statement of the above global conditions on island constraints does make the implicit claim that these conditions form a natural class.

In Chapter Four, I will argue that the phenomena described in this section by global conditions follow in a natural way from independently supported assumptions made by the Behavioral Position.

1.2.7. Before concluding our review of Ross' dissertation, it is worthwhile to look at the most important questions Ross considered his thesis had raised without answering:

(1.99) a. Why should rules which adjoin terms to the right side of a variable be upward bounded, and not those which adjoin terms to be the left of a variable?
b. Why should it be that chopping rules, feature-changing rules and unidirectional deletion rules share the property of being subject to the constraints, to the exclusion of other rules?
c. Why should there be a difference between unidirectional and bidirectional pronominalization?
d. Why should complex NP's, coordinate nodes, sentential subject clauses and NP's on the left branches of larger NP's all function the same in defining islands?
e. Can islands be shown to behave like psycholinguistic entities?

A large part of the remainder of this thesis will be devoted to arguing that the answer to (1.99e) is yes. With respect to (1.99a)-(1.99d), I believe they rest on mistaken assumptions and should be reformulated before they can be answered. I will discuss (1.99a) in 2.1.3 and show that the pertinent constraint is much weaker than Ross though. Concerning (1.99b) and (1.99c), I have already shown that unidirectional and bidirectional pronominalization are irrelevant notions with respect to island constraints; moreover, I have also pointed out that the notion 'feature-changing rule' should probably be rephrased as 'neutralization', to cover at least the comparator-comparee relation, which has nothing to do with feature-changing. Finally, I shall argue in Chapter Three and in Chapter Four that only complex NP's, coordinate nodes, and sentences or verb phrases modified by optional adverbials function the same in defining islands, as they are the only configurations the constraints on which are descriptively collapsible and therefore capable of receiving the same explanation. Once Ross' questions have been appropriately reformulated, the answer to them will essentially consist of the answer to (1.99f).

(1.99) f. How do islands behave as psycholinguistic entities?

1.3. There are two more proposals within the Transformational Position which we shall consider in this chapter. The first represents a reaction of Chomsky's to Ross' criticism of the A-over-A principle; Chomsky attempts to defend his principle within the framework of his Extended Standard Theory. The somewhat more elegant formulation of the principle given in this paper is reproduced below as (1.100), although the latter is, as far as I can see, a notational variant of (1.4).

(1.100) If a transformation applies to a structure of the form
$[_{\alpha} \ldots [_{\underline{A}} \ldots] \ldots]$
where α is a recursive node, then it must be so interpreted as to apply to the maximal phrase of the type A.

1.3.1. Chomsky writes that "it would be a welcome result" if a "careful formulation of the A-over-A condition", together with a set of conditions to be examined below, would turn out to be capable of adequately handling the range of examples earlier treated as conditions on transformations, for the A-over-A principle seems to be an intuitively natural one, while "such conditions as the Complex Noun Phrase Constraint seem quite ad hoc". And indeed, it is not difficult to assent to the claim that if Ross' treatment represented a step forward over the A-over-A principle in terms of observational adequacy, it looked very much like a step backwards in terms of descriptive adequacy. However, as no account can be descriptively adequate without also being observationally adequate, it will be interesting to consider in some detail the ways in which Chomsky proposes to make his analysis observationally adequate.

Earlier, I mentioned two kinds of cases in which the A-over-A principle turned out to be too strong; the two kinds were illustrated by the examples (1.6) and (1.7), (1.10), (1.13) respectively, which I reproduce below for convenience:

(1.6) You approve of my seeing someone.

(1.7) You saw the picture of the master of the house.

(1.10) I believe that John claimed that Mary insinuated that Bill was guilty of murder.

(1.13) John decided to try to begin to write a book.

In connection with (1.6), where the NP *someone* can be wrenched out of the containing NP *my seeing someone*, Chomsky has in fact no solution. He proposes to weaken the A-over-A principle from the status of a universal constraint to that of a part of an evaluation measure on the complexity of grammars. Thus, *who would you approve of my seeing* is merely a 'marked' form, rather than an ungrammatical one, since the violation of the A-over-A principle does no longer necessarily result in ungrammaticality. What this account leaves unexplained is why the above marked form should be perfectly acceptable, while other instances of violation of the principle should be totally unacceptable.

In connection with (1.7), (1.10) and (1.13), where any of the lower NP's, S's, or VP's can be fronted, Chomsky claims that the A-over-A principle is not in fact involved. His explanation runs as follows: there are, among other rules, the cyclic rules of *wh*-placement and of *wh*-movement; the rule of *wh*-movement refers only to constituents that carry a *wh* (in Chomsky's system, topicalized elements also carry a *wh*-like feature); therefore, the chopping rule of *wh*-movement is unambiguous, and the A-over-A principle does not need to be invoked. However, as Chomsky nowhere says that the A-over-A principle constrains only chopping transformations, it is not at all clear why it allows the placement of *wh* on the lower A's. *Wh*-placement may be regarded as a feature-changing transformation, but I think it is more correctly described as the overt manifestation of a neutralization phenomenon, of the type discussed already; in the case of relative *wh*-words, the neutralization involves conferentiality, in the case of interrogative *wh*-words, it involves the 'target' of a question-verb (in the performative analysis, this description covers direct questions as well). Such neutralizations are subject to island constraints, as we saw above, and *wh*-neutralization is no exception. To see that, we must consider a situation in which *wh*-words cannot move; this arises when there are more *wh*-interrogatives than question verbs, as in (1.101).

(1.101) a. Someone went to Paris and to some other place.
b. *Who went to Paris and to which other place?

The fact that (1.101b) cannot be derived from (1.101a) indicates that *wh*-neutralization must be constrained by the A-over-A Principle or some equivalent device.

1.3.2. By denying that *wh*-neutralization is subject to the A-over-A Principle, Chomsky must impose a whole new list of conditions on the rules which move *wh*-like words; the applicability of the A-over-A Principle is, of course, severely restricted by this decision. On the other hand, the A-over-A Principle becomes applicable in a new range of cases, which did not require any meta-grammatical constraints in Ross' framework, due to some rather radical changes in Chomsky's theoretical framework. For example, due to the elimination of the Subject-Raising rule and to the definition of the structural index of the passive in terms of categories rather than of relations, the A-over-A condition becomes important in preventing ungrammatical outputs of the Passive transformation.

In order to block the movement of one term of a coordinate node, Chomsky resorts to a convention proposed by Dougherty to the effect that feature complexes are not assigned to coordinate terms directly, but rather to the coordinate node first, after which they 'percolate' into the coordinate terms. Since a feature like *wh* has to be assigned both to coordinate NP-node and to the NP-nodes that it immediately dominates, the A-over-A principle will effectively block the movement of one of the latter. It is not clear, however, how the movement of an element of a coordinate term will be blocked. If the coordinate node is an S, it will be rather artificial to assign it a *wh*, in order that some NP inside the coordinate S's be eventually marked with this feature. Moreover, it will be necessary to allow the 'across-the-board' movement of elements in violation of the A-over-A principle.

It is not very clear how the movement of the head of a relative clause can be prevented, so long as neutralization manifested by *wh*-placement is not subject to the A-over-A condition. Presumably, the percolation-convention would have to be extended to complex NP's, requiring that the head can only receive a *wh* if the higher NP node received one as well.

With respect to the Left Branch Condition, Chomsky suggests in note 10 that it be made part of the larger condition which blocks the movement of 'specifiers' of NP's and AP's (determiners and pre-adjectival adverbial modifiers respectively). This more general formulation, as well as the separation of the freezing of specifiers and of Complex-NP-heads (which were both collapsed under the LBC in Ross' treatment), seems to me essentially correct; in fact, I will argue that it was wrong to blame the freezing of complex NP heads on the LBC.

1.3.3. A large number of restrictions are taken care of by the following rather complex constraint:

(1.102) No rule can involve *X*, *Y* in

...*X*...[α...*Z*...-*WYV*...]...

where (i) *Z* is the subject of *WYV* and is not controlled by a category containing X
or (ii) α is a subject phrase properly containing *Y*
or (iii) *Y* is in COMP and *X* is not in COMP
or (iv) *Y* is not in COMP and α is a tensed S
or (v) *X* is in COMP and *Y* is not in an adjacent cycle.

1.3.3.1. (1.102i) concerns a number of cases which were not discussed in Ross' dissertation, some of which, Chomsky claims, have the ultimate effect of reducing ambiguity in language. These cases will be briefly considered in Chapter Two.

1.3.3.2. (1.102ii) is in effect equivalent to Ross' Sentential Subject Constraint, except that it does not allow for the optionality of Pied-Piping in subject position, and thus rules out the grammatical (1.103b).

(1.103) a. A picture of John hangs in the dining-room.
b. It's John who a picture of hangs in the dining-room.

1.3.3.3. (1.102iii) presupposes Bresnan's Complementizer Substitution Universal. In this view, all sentences in all languages are generated in the base with a node COMP[11] (for complementizer), which is necessarily lexically empty. This node may appear to the right or to the left of sentences, and it is only in the latter case that transformations can insert lexical material into COMP. The lexical material that can be so inserted consists by and large of wh-like words (i.e., relativized, questioned or topicalized constituents in various languages). The condition that COMP-substitution transformations cannot operate if the COMP is clause-final is in fact equivalent to Ross' Right Roof Constraint. If COMP is not filled by a wh-phrase, it is optionally filled by that (incidentally, Dean (1967) pointed out that verbs like quip, croak, roar, etc., require a that complementizer). What (1.102iii) does is ensure that the rules that Ross views as unbounded and that Chomsky regards as proceeding by cyclic 'hopping' will only adjoin a chopped constituent to the left of sentences.

1.3.3.4. (1.102iv) prevents elements from escaping from tensed interrogative clauses. Thus, in both (1.104) and (1.105) Y (in these instances, what) cannot be in COMP, for the COMP position is filled by where and how respectively, and the movement of what is blocked on the inner cycle. On the outer cycle, however, what can escape in (1.105), but not in (1.104), for, in the latter, it originates in a tensed clause.

(1.104) *What did he wonder where John put?

(1.105) What crimes does the FBI know how to solve?

In notes (25) and (26), Chomsky admits that (1.102iv) is too strong for many speakers in blocking movements out of all tensed interrogative clauses (some speakers accept What crimes did he wonder how they solved) and that (1.102i) is also too strong, as it stars (1.105).

From the way in which (1.102) is formulated, it is not clear whether or not the set (i-v) is intended as an exhaustive list of conditions; that is, it is not clear whether the where that precedes (1.102i) is to be interpreted as an if and only if or merely as an if. Given the fact that (1.102) is the last of a number of reformulations, we can assume that an if and only if-statement was attempted, and infer from (1.102ii) and (1.102iv) that if X is in COMP and Y is in COMP, the result will be grammatical. In this connection, I should like to point out a rather interesting counter-example from Rumanian.

---

[11] Notice that this proposal comes very close to a tacit acceptance of the performative analysis, since it is hard to see what other interpretation could be given to the fact that main clauses are introduced by COMP.

In Rumanian, it is possible to move more than one questioned constituent in COMP position in the same clause. Thus, the following sentences are all grammatical:

(1.106) Cine pe cine a întînit?
Who whom met?
"Who met whom?"

(1.107) Nu ştiu cine unde a plecat.
I don't know who where went.
"I don't know who went where."

(1.108) Ti-am spus deja cine pe cine cînd a ucis.
I have told you already who whom when killed.
"I have already told you who killed whom when."

Apparently, the base rules of Rumanian generate more than one clause-initial COMP-node. One might attempt to maintain the one-COMP-per-S hypothesis by proposing that the various 'wh-words' are in effect coordinated and that the coordinate node is inserted under the COMP-node. Such a hypothesis would be, however, disproved by (1.109) and (1.110).

(1.109) *Nu ştiu cine cînd unde a murit.
I don't know who where when died.
"I don't know who died where when."

(1.110) Nu ştiu cine unde şi cînd a murit.
I don't know who where and when died.

Apparently, the COMP-nodes of Rumanian have to be marked for the category they can accommodate, and there is only one COMP for adverbials. Thus, (1.110) is grammatical, for only one (coordinate) adverbial node has been inserted, while (1.109) is ungrammatical, because one of the adverbs has failed to find an empty COMP-node. This strongly suggests that there is more than one COMP-node in (1.106)-(1.108), and that the various questioned constituents in these sentences are not coordinated.

Another fact which suggests that the equivalents of wh-words in (1.106)-(1.108) are not coordinated is the existence of the following minimal pair:[12]

---

[12] I am indebted to my wife, Mariana Grosu, for this observation.

(1.111) a. Nu ştiu *cine pe cine* a prins.
b. Nu ştiu *cine, şi pe cine*, a prins.

In (1.111b), the wh-words are overtly coordinated, and the a- and b-sentences are not quite identical in meaning. (1.111a) translates roughly as (1.112a), while (1.111b) translates roughly as (1.112b); in the a-sentences, the speaker asserts that he knows one fact, while in the b-sentences he asserts that he knows two facts.

(1.112) a. I don't know who caught whom.
b. I don't know who$_i$ caught someone$_j$, and I don't know whom$_j$ he$_i$ caught.

As pointed out above, it would follow from Chomsky's principle that a wh-phrase that has moved into COMP-position can move upwards. But this is not the case for Rumanian, as the movement of any of the wh-words in (1.108) results in total ungrammaticality; this can be seen in (1.113)-(1.115) (the last is grammatical, but only if *cînd* 'when' originates in the matrix sentence).

(1.113) *Cine ţi-am spus pe cine cînd a ucis?
Who did I tell you whom when killed?

(1.114) *Pe cine ţi-am spus cine cînd a ucis?
Whom did I tell you who when killed?

(1.115) Cînd ţi-am spus cine pe cine a ucis?
When did I tell you who whom killed?

What these facts suggest is that the explanation of the restrictions on chopping out of interrogative clauses in terms of wh-hopping is mistaken. In addition, it still remains to give an observationally adequate account of those cases in which interrogative clauses function as islands, since neither Chomsky nor Ross have done anything of the kind. Perhaps, the solution to this difficult problem should be sought not only in formal features, like the presence or absence of tense in the interrogative clause, but also in the semantics of each interrogative word. The matter remains open, for the time being.

Since Chomsky requires that COMP dominate a null lexical string in the base, I assume that he would not want to claim that subordinating conjunctions like *although*, *unless*, *therefore*, etc., occupy COMP-position at any time. Notice, however, that if such conjunctions are construed as complementizers, and if the theory is modified to allow their generation in the base in such a position, then (1.102iv) is powerful enough to account for what I called earlier the Optional Adverbial Constraint. Additional specifications would, however, be necessary, since adverbials seem to disallow the migration of their elements even when they are not tensed, and even when they are not clauses.

1.3.3.5. Finally, (1.102v) is meant to cover some cases which fall under the Complex NP Constraint. Indeed, although both S and NP are cyclic nodes, it is S alone that has a COMP, and some wh-word inside a clause in apposition to a sentential noun can only move in the COMP of that clause; it will be unable to move on the next cycle, since the higher NP phrase has no COMP, and (1.102v) prevents Y from 'skipping a cycle' (this condition is unnecessary to prevent movements out of relative clauses, since those have a wh-word in COMP position, and chopping is thus blocked one cycle earlier).

If (1.102v) is adequate for complex NP's, it fails again by being too strong in the case of embedded interrogatives. Indeed, it rules out the entirely acceptable (1.116).

(1.116) This is the book which I am not sure whether (or not) I told you how to bind.

This additional example strengthens the case for not attempting to handle interrogative-clause islands in purely formal terms.

1.3.4. In conclusion, it appears that Chomsky's analysis of the phenomena considered by Ross does not go very far in providing a natural alternative. Chomsky rightly criticizes Ross for proposing a list of ad hoc constraints, but it is difficult to see why his (1.102ii)-(1.102v) chould not be subject to the same criticism. Also, the A-over-A Principle, which did seem to have an air of naturalness about itself in Chomsky's initial treatment, does not live up to expectations in Chomsky (1971), owing to the large number of cases in which it makes wrong predictions, or in which it makes correct predictions for the wrong reasons (e.g., the blocking of movements of elements of a coordinate term just in case the element to be moved and the coordinate node happen to belong to the same category). I believe we must conclude that Chomsky's 1971 proposal comes no closer than Ross' 1967 one to providing a natural solution to the phenomena at issue.

1.4. The last attempt to arrive at a descriptively and explanatorily more adequate account of island constraints within the Transformational Position which we shall examine in this Chapter is Ross' Island Constraint (or the IC), orally proposed in a series of lectures at The Ohio State University (spring, 1971).

(1.117) The IC: If some constituent X of category A immediately dominates some constituent Y, also of category A, then no constituent Z of any category type whatever, which is dominated (not necessarily immediately) by X can be moved from under the domination of X.

The IC revives the A-over-A Principle, and attempts to avoid its observational inadequacies by appropriately weakening and strengthening it; this is done by the first and second underscored specifications respectively.

Although I consider the specification of immediate domination to be a valuable insight (which I shall incorporate into my counter proposal of Chapter Three), I think that the IC must be rejected because of at least the following observational deficiencies:

(1.118) a. It wrongly predicts that clauses moved by Extraposition (and attached to the first S-node up) are islands, as can be seen in (1.119).

(1.119) Who is it possible that Mary loves?

(1.118) b. It wrongly predicts that Extraposition-from-NP is an impossible rule.
c. It wrongly predicts that Extraposition-of-PP is an impossible rule.
d. If prepositional pharses are regarded as NP's, it wrongly predicts that prepositions cannot be stranded.
3. It wrongly predicts that nominal complements of NP's cannot be moved, that is, that (1.120a) and (1.120b) are ungrammatical.

(1.120) a. Which house did I show you a picture of?
b. Of which cars were the hoods damaged by the explosion?

(1.118) f. If complementizers are Chomsky-adjoined to the clauses they introduce, it wrongly predicts that (1.121) is ungrammatical.

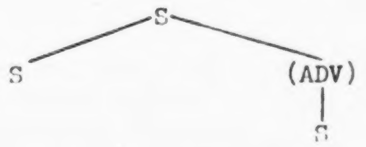
(1.121) Who did you claim that John saw?

(1.118) g. It wrongly predicts that elements of a main clause modified by an adverbial clause cannot move, that is, that (1.122) is ungrammatical.

(1.122) The woman who John dislikes, because he suspects her of deceit, is seriously ill.

The applicability of the IC to such structures follows from the claim made in Ross (1967) that they have essentially the surface structure in (1.123).

(1.123)



Ross defended the structure in (1.123) with two arguments: (i) the natural place to pause in (1.124) is before because, which suggests that this is where the major structural boundary lies.

(1.124) John hit Bill because he was furious.

(ii) Extraposition cannot move a clause around the adverbial, as seen in (1.125), a fact which would follow from a structure like (1.123) and the RRC.

(1.125) a. That Mary is sick is funny because no one has been sick for years.
b. It is funny that Mary is sick because no one has been sick for years.
c. *It is funny because no one has been sick for years that Mary is sick.

The two above arguments are not entirely conclusive, for the major break in a phrase does not necessarily allow pause (e.g., one does not usually pause between subject and verb phrase), and the facts of Extraposition do not work in the same way for non-clausal adverbs (the reason is, I believe, that the RRC is an incorrect principle, as I shall argue in 2.1.3), as can be seen in (1.126).

(1.126) a. That John was sick was announced yesterday at noon.
b. It was announced that John was sick yesterday at noon.
c. It was announced yesterday at noon that John was sick.

As it is rather implausible that adverbs are differently adjoined to higher nodes depending on whether they are clausal or not, Ross' second argument is inconclusive. But there are better arguments which can be adduced to support his position, Thus, (iii) As two adverbs of the same kind cannot occur in the same clause, we must assume that the second adverbial in (1.127a) and (1.127b) belongs to a higher clause than the main one.

(1.127) a. John struck his kid because he had been naughty, because Mary asked him to do so.
b. John went out to buy some whiskey, to please his fiancee.

(iv) Negative polarity items like any or ever are tolerable inside adverbs, provided that the latter are not preceded by pause, as seen in (1.128).

(1.128) a. John wasn't arrested because he had ever killed anyone (he was nabbed for strong-armed robbery).
b. *John wan't arrested, because he had ever killed anyone.

The difference in acceptability between (1.128a) and (1.128b) follows quite naturally if we assume that the latter only has the structure in (1.123), since the negative polarity items would be commanded by negation in (1.128a), but not in (1.128b).

I regard the arguments (iii) and (iv) above as sufficient for concluding that sentence-modifying adverbials have the structure in (1.123), and that (1.122) is a counterexample to the IC.

(1.118) h. The IC wrongly predicts that adverbs cannot be clefted, and this is not in general true, as shown in (1.129).

(1.129) a. It's because he loves her that John beats his wife.
b. It was (in order) to gain Mary's confidence that John pretended not know that she was rich.
c. It's since he began the research on his dissertation that Bill has gone nuts.
d. It's after the battle that scores are settled.
e. It's before he arrived that she died.[13]

---

[13] Many other adverbial types can be clefted, such as the ones introduced by instrumental with, before, after, temporal since, etc. However, although, as, for, causative since, and perhaps others, cannot be clefted; the latter three are interesting in that they seem to be logically synonymous with because, but behave differently with respect to clefting, as (i) shows.

(i) a. It's because he had no money that John dropped out of school.
b. *It's $\left\{\begin{array}{l}\text{as}\\ \text{for}\\ \text{since}\end{array}\right\}$ he had no money that John dropped out of school.

M. Geis pointed out to me that the adverbs in (ib), unlike the one in (ia), disallow the chopping of elements of the main clause--as seen in (ii)--, which constitutes an apparent counterexample to my claim (3.2.ii) in Chapter Three, to the effect that movements of elements of the main clause are free in principle.

(ii) a. Who did John jilt because she had no money?
b. *Who did John jilt $\left\{\begin{array}{l}\text{as}\\ \text{for}\\ \text{since}\end{array}\right\}$ she had no money?

A third distinguishing property is the ability to occur in embedded clauses, as shown in (iii).

(iii) a. I believe [John left the country because he was unhappy].
b. *I believe [John left the country $\left\{\begin{array}{l}\text{as}\\ \text{for}\\ \text{since}\end{array}\right\}$ he was unhappy].

A fourth distinguishing property is ambiguity under negation. Thus, (iva) is ambiguous, as pointed out in Lakoff (1970), while (ivb) is not.

(iv) a. John doesn't beat his wife because he loves her.

b. John doesn't beat his wife $\left\{\begin{array}{l}\text{as}\\ \text{for}\\ \text{since}\end{array}\right\}$ he loves her.

A fifth distinguishing property was pointed out to me by J. Hutcheson: while (va) can be a well-formed answer to *why does John beat his wife?*, (vb) cannot.

(v) a. Because he loves her.

b. *$\left\{\begin{array}{l}\text{As}\\ \text{For}\\ \text{Since}\end{array}\right\}$ he loves her.

A complete study of adverbs is beyond the scope of this dissertation, but it seems to me that the data presented so far suggest that *as*-, *for*-, and *since*-, clauses *must be comments of the speaker*, while this is not necessary for *because*-clauses. If this is correct, we have a fairly straightforward explanation for the five facts above.

Thus, (iiib) is bad, because a comment of the speaker must be directly dominated by the top S-node.

(iva) is ambiguous because the *because*-clause can be dominated either by the top S, in which case it commands the negation, or by the VP-node of the main clause, in which case negation commands it; as the latter situation could not arise for *as*-, *for*-, or *since*-clauses, (ivb) is unambiguous.

With respect to the unacceptability of (iib), a number of writers have proposed that unbounded movement rules, such as Question-Movement, attach an element through Chomsky-adjunction (i.e., by creating a superordinate node of the same kind as the one to which the element is attached) rather than by sister-adjunction. Very little evidence has been given in favor of this position, as far as I know, but if *who* is indeed Chomsky-adjoined to the top S-node, the unacceptability of (iib) will follow from the fact that the adverbial clauses are no longer dominated by the top S-node; thus, the badness of (iib), supported by the conclusions that we drew from the badness of (iib) and (ivb), may be regarded as evidence that unbounded movement rules perform Chomsky-adjunction.

The badness of (ib) and (vb) are, I believe, different manifestations of the same fact. Thus, it has often been pointed out that clefted (and pseudo-clefted) constituents, as well as (appropriate) answers to questions, represent the *main* assertion of a sentence, or, as certain writers put it, the new information which a sentence conveys; this is also mirrored in the fact that cleft and pseudo-cleft constituents as well as answers to questions (and question-words, for that matter), must always bear stress. But *as*-, *for*-, and *since*-, clauses constitute (parenthetical) comments of the speaker on the proposition expressed in the main clause, and consequently *cannot* be the *main* assertion of the entire sentence; hence, they cannot be cleft constituents or answers to questions, and of course, there can be no question-word applicable to such constituents (that such clauses cannot be pseudo-clefted is shown in (vi), where the a-sentence is possible for some speakers, while the b-one is out for everyone).

(vi) a. Why I didn't want to see you was because you had always been nasty.

b. *Why I didn't want to see you was $\left\{\begin{array}{l}\text{as}\\ \text{for}\\ \text{since}\end{array}\right\}$ you had always been nasty.

The import of the facts discussed in this note is that very slight semantic distinctions can have notable syntactic consequences. The same was argued in note 3 to this chapter, where it was pointed out that *exceed* and *more than*, although apparently logically equivalent, cannot have the same semantic representation. Such cases undoubtedly deserve further study (for a discussion of related problems, see Shopen (1972)).

(1.118) i. The IC falsely predicts that adverbials cannot be preposed from under the dominance of the immediately higher S-node (the circled one in (1.123), and this is contradicted by (1.130b), which is derived from (1.130a).

(1.130) a. I wonder whether I could talk to you, in case you have a spare minute.

b. In case you have a spare minute, I wonder whether I could talk to you.

(1.118) j. The IC falsely predicts that all coordinate structures are frozen to the same degree. In Ross (1967), the weaker claim that only symmetric coordination is subject to the CSC is made, but it will be shown in detail in Chapter Three that the situation is much more complicated than that; in fact, different types of coordination are subject to different constraints. Ross failed to consider all the pertinent cases, and even the ones he considered are not constrained exactly as he thought.

k. The IC (as well as the CSC) falsely predicts that across-the-board movements out of coordinations cannot occur, that is, that (1.131) is ungrammatical.

(1.131) It's potatoes that Bill likes and Jill dislikes.

I take it that the objections (1.118a)-(1.118k) are sufficient to show that the IC is not a tenable proposal, even though I think it comes closer to descriptive adequacy than the other alternatives we have considered so far.

In Chapter Three, I shall present a different proposal which, as far as I can see, avoids the counterexamples that viciate the IC, and lends itself to a natural explanation. My criticism of the various proposals I have discussed was not meant to belittle them. On the contrary, I regard them as highly interesting attempts to make contributions to the study of syntax, and the fact that empirical proposals exhibit flaws should come as no surprise. In fact, I have no doubt that my own proposal, even if ranking higher at various levels of adequacy than those I have rejected, will itself turn out to be inadequate in a number of ways.

CHAPTER II

PERCEPTUAL PRINCIPLES AND GRAMMAR

2.0.0. This chapter is essentially concerned with providing linguistic illustrations of the applicability of the three kinds of perceptual principles mentioned in the Introduction, namely, the principles of erroneous closure, the principles of perceptual conflict, and the principles of interrupted behavior; these three kinds of principles will be discussed in sections 2.1.1, 2.1.2, and 2.1.3 respectively (the last will be introduced in 2.1.2).

The theory of perceptual strategies summarized in Bever (1970) is briefly outlined in 2.0.1. The last major section of this chapter, 2.2, is concerned with various possible minimal distance principles, and their failure (as far as I have been able to ascertain) to affect grammars in significant ways.

2.0.1. In earlier writings within the framework of transformational grammar, the distinction between "ungrammatical" sentences and sentences "unacceptable for performance reasons" was presented as a hard and fast one.[1] Most recently, various writers have questioned the validity of this absolute distinction. In particular, T. G. Bever has claimed, in a number of articles, that the internal mechanisms of speech production, of speech perception, and of prediction of the acceptability of potential sentences are partially independent and mutually influence one another. The part of his hypothesis that we shall be essentially concerned with is the claim that the strategies employed in speech perception may constrain the acceptability, and sometimes even the grammaticality, of sentences, and that the outright rejection of certain constructions by native informants may be explained in a much more illuminating way by considering possible perceptual causes than by seeking a purely formal analysis.[2]

---

[1] There has also been a great deal of discussion on the treatment of semantic ill-formedness, i.e., whether the latter should be handled as a syntactic or extrasyntactic phenomenon, but this is largely irrelevant to the issues we shall be concerned with.

[2] I wish to stess that my adoption--as a working hypothesis--of Bever's claim that features of perception may affect the acceptability of sentences should not be construed as an acceptance of every aspect of his theory of interacting systems.

In particular, I disagree with his statement that "transformations are irrelevant to most ongoing speech behavior." This statement rests on the fairly well supported notion that the production of a sentence does not simulate its generation by a transformational grammar, and that its understanding does not simulate a reversal of the grammatically pertinent rules. The most obvious argument against equating grammatical rules with real-time processes is that the rules of grammar operate on the entire string, while a sentence begins to be interpreted as it begins to be heard and need not exist in its entirety in the head of the speaker before it is pronounced.

Notice, however, that this only shows that transformations are probably not used as real-time processes, but not that they are not used at all in behavior. In fact, I find it hard to conceive of a person producing or understanding certain sentences without using his knowledge about transformational relations, or about other features of his language. Indeed, how could one say or understand

(i) although he is poor, John married a rich girl, or
(ii) Bill killed his wife and was in turn killed by his mistress

without consciously knowing that poor and rich are antonyms, or that the active and the passive are converse relations? Conversely, how could one construct a pun without being aware of ambiguities? It seems to me that everything a person knows about his language may be actively utilized in production or perception, even though this utilization need not take the form of grammar-simulation.

Bever's statement that "competence is performance" has drawn sharp criticism. On the face of it, this statement is not just hard to believe, but also in flagrant contradiction with Bever's other claim that transformations are intuitions not used in actual speech behavior, which implies that one's knowledge about language and one's linguistic behavior are at least partially distinct. I believe, however, that these difficulties are due to Bever's failure to define his notion of competence carefully enough. Indeed, Bever uses "competence" in the special sense of "what a speaker consciously knows about his language," and claims that a speaker's intuitions are limited to acceptability judgments, but shed no light on the sources of unacceptabilities. Thus, if competence is defined as the sum of a speaker's intuitions and if expressing intuitions is a form if performance, it is clear that the distinction between performance and competence becomes a trivial one, and one may claim that competence is part of performance, or that "competence is performance" (which does not mean that performance is competence).

Bever refers to a number of experiments which suggest that the various mechanisms which are integrated in the adult are acquired independently by the child. Thus, children seem to go through three stages during their linguistic development. Very young children appear to depend on "basic linguistic capacities," such as the belief that words refer to objects and actions in the world, or that there are functional internal relations between parts of sentences, such that they express the roles of the various actants. Children around the age of seven begin to be able to appreciate the regularity of the relation between actives and passives, which suggests that they have begun to acquire a grammar, i.e., an abstract system which enables them to make sophisticated predictive judgments about sentences. However, children between the ages of two and six appear to rely rather heavily on perceptual strategies like (2.1) and (2.2).[3]

(2.1) The first N...V...(N)...clause is the main clause, unless the verb is marked as subordinate.

(2.2) Any Noun-Verb-Noun (NVN) sequence within a potential internal unit in the surface structure corresponds to "actor-action-object."

Thus, in an experiment conducted with children between 1 1/2 and 2 1/2, it was found that when presented with a complex sentence like the cow that jumped walked away, and asked to act out, some children act out the doubly-underscored portion. It appears that they considered the first clause to be the main clause, in other words, that they followed (2.1) completely, and that they were unable to make use of overt marks of subordination. It is hard to tell whether they can recognize such marks at all, but whatever the situation, a strategy like (2.1) appears to override other considerations. Similarly, it was found in another experiment that the performance of children on "reversible" passives (i.e., sentences in which semantic or contextual information do not give any clue as to the roles of the actants) deteriorates considerably around the age of four. The suggested conclusion is that children begin to depend on a strategy like (2.2) at that age.

Bever does not attempt to decide whether strategies like (2.1) and (2.2) are inductions over experience or whether they are determined by maturational development. He points out, however, that even an empiricist view of the acquisition of such strategies which would claim that they "are formed in response to natural probabilities in the actual speech that the child experiences," would have to explain why the children select precisely these strategies out of the many alternative possible generalizations; in other words, it would have to "include a nativist component" (Bever, 1970).

The claim that children acquire a number of perceptual strategies before they have a complete grammar is crucial for the claim that grammar may be constrained by perception. Indeed, if certain perceptual strategies already exist at the stage at which grammatical generalizations are acquired, the child may constrain the form of some of his grammatical rules by incorporating perceptual restrictions directly into his predictive system. In such a case, we can say that a behavioral phenomenon has been "grammatized." This is, however, not the only possibility, for the grammar the child acquires need not be, and in all probability never is, affected by *all* his perceptual strategies. In a large number of situations, perceptual strategies survive without affecting the grammar, and they lead to unacceptabilities of various degrees, rather than to downright ungrammaticality.

---

[3] These are, respectively, the principles Bever labels (B) and (D) in Bever (1970).

In addition to the "positive" strategies that children acquire, there are almost certainly constraints as to what kinds of strategies can be acquired. To the extent that such constraints reflect limitations of the processing device, they must be universal. However, as behavioral limitations determine gradual, rather than abrupt, complexity, they can lead to two kinds of situations, exactly like the positive strategies: they can either be grammatized, and thus become grammatical constraints, or they can survive as purely behavioral limitations, in which case the degree of unacceptability will be proportional to the "amount of complexity" present in the string.

As Bever himself points out, it is not at all a simple matter to decide which cases of unacceptability are due to "purely syntactic" factors and which are due to behavioral ones. Indeed, the only data available to us are intuitive judgments as to the relative acceptability of discourse, but intuition is not a very reliable tool for discriminating between ungrammatical and "strictly unacceptable" sentences, and is in general powerless to ferret out possible behavioral causes when some unacceptability has been grammatized. Therefore, it is to be expected that the causes of certain unacceptabilities be undecidable, at least for the time being, and that even the most believable claims as to perception affecting grammar may not be provable beyond a certain degree of doubt. I believe, however, that we can regard a case as having been convincingly presented when it can be argued that some set of apparently unrelated phenomena (which could be accounted for formally only by invoking arbitrary and ad-hoc features) receive a simple and straightforward explanation in behavioral terms.

With respect to the problem of distinguishing ungrammatical from merely unacceptable sentences, it would be unreasonable as I pointed out above, to require a fool-proof procedure. However, some rules-of-thumb may prove useful.

With respect to erroneous closure, whose effect is to "lead the hearer up the garden path" (so that the acoustic signal has to be *reinterpreted* if one is to make sense out of it), we may regard the difficulty as grammatized, if the sentence remains bad *even after it has been pointed out to the hearer that he had been led up the garden path*; on the other hand, if the sentence becomes acceptable when the hearer has been told where closure should take place, we may assume that there has been no grammatization.

Concerning cases involving interrupted behavior, we may assume that grammatization has not occurred, if we get a *gradual acceptability scale* (rather than a *good/bad dichotomy*), as we vary the values of the pertinent parameters (length, complexity, predictability, etc., of the intervening material). An apparent counterexample to this generalization would be the case of multiple center-embeddings, where there seems to be an abrupt jump from tolerable acceptability to very high unacceptability when we move from one degree to two degrees of embedding. However, center-embeddings are essentially different from the cases in which we get gradual judgments. Thus, the latter are found with *one* instance of interrupted behavior, while center-embeddings to a degree higher than one involve at least *two* independent instances of interrupted behavior which must be handled simultaneously; the situation is therefore radically different, and, in addition, there seems to be some perceptual conflict involved in multiple center-embeddings (Bever, 1970; for additional discussion, see notes 18 and 19 to this chapter).

Regarding cases which exhibit perceptual conflict, we may assume that grammatization has not occurred if sentences become acceptable after we have trained ourselves to overcome the difficulty involved and to puzzle out the meaning of the corresponding sentences; if the sentences are still bad when their meaning can be apprehended with comparative ease, we may regard this as an indication of grammatization. For example, if perceptual conflict is involved in *John's father's brother* or in *it's not the case that John's not going to the party had no effect on Mary*--as claimed in Bever (1970) (see also note 18 to this chapter)--I believe that the conflict is not grammatized; in particular, expressions of the former type (and of the same or greater length) are probably used quite commonly in anthropological literature. On the other hand, if perceptual conflict is involved in the badness of *he$_i$ said that John$_i$ was sick*--as I argue in Chapter Four--grammatization probably *has* occurred, for the feeling of deviance persists after it has been made clear to the hearer that *he* and *John* stand for the same referent.

By the criterion I have just offered, center-embeddings like *the boy the teacher the girl kissed spat at died* would be regarded as ungrammatical rather than merely difficult, for it is hard to believe that people could become entirely comfortable with such sentences. Most writers have, however, argued that the grammatical restriction would be hard, if not impossible, to state, as it is not true that all center-embeddings are as bad as the previous one; for example, *the surgeon who the girl*!#*that hoodlum raped was afraid of # had won high distinctions in medical school* is quite acceptable (the symbol # indicates overlong pauses). Therefore, it is at least doubtful whether multiple center-embeddings should be regarded as ungrammatical in English. This indicates that the various criteria I have proposed do not always yield the desired results; this is, however, not surprising, since I have made it clear that they are rules-of-thumb, and nothing more.

2.1. We shall now consider the three perceptual principles mentioned above in some detail.

2.1.1. The strategy (2.1) mentioned in 2.0.1. is one factor capable of giving rise to erroneous closure. The point is discussed by Bever (1970), who acknowledges his indebtedness to Kirk. Thus, as can be seen from the paradigm in (2.3), when a subordinate clause appears initially, the mark of subordination is not deletable, even though it may be deleted in other positions.

(2.3) a. It is odd that John dislikes sweets.
b. It is odd John dislikes sweets.
c. That John dislikes sweets is odd.
d. *John dislikes sweets is odd.

Bever also points out that strategy (2.1) predicts that subject relative pronouns of restrictive relative clauses are not deletable, as the following tensed verb would be interpreted as the verb of the main clause. However, if the tensed verb is _be_, and if the latter is in turn deleted, the result is grammatical, for either the remainder of the verb is not tensed and can only be subordinate, or there is no verbal form left which would enable (2.1) to interpret the string erroneously. These facts are exhibited in (2.4).

(2.4) a. The man who came late forgot his money.
b. *_The man came late_ forgot his money.
c. The senator who was in the saddle shot from the hip.
d. *_The senator was in the saddle_ shot from the hip.
e. The senator in the saddle shot from the hip.
f. The monkey who was running after the bus slipped on a banana.
g. *_The monkey was running after the bus_ slipped on a banana.
h. The monkey running after the bus slipped on a banana.

The underlined sequence in (2.4b, d and g) is interpreted by (2.1) as the main clause, and the remainder of the sentence in question cannot be made sense of.

As far as I can tell, (2.1) is grammatized in most (possibly in all) dialects of English with respect to the facts exhibited in (2.3) and (2.4); that is, there are no dialects that I know of in which (2.3d), (2.4b), (2.4d) or (2.4g) are grammatical. There is, however, an interesting difference between (2.3) and (2.4): while a complementizer can be deleted in object position--as shown by the grammaticality of (2.2b), the deletion of a subject relative pronoun is usually prohibited in both subject and object position, even though this does not follow from (2.1); that is, many dialects do not allow _I'm a man loves his beer cold_, although some do (the sentence seems to be at least easier to comprehend than (2.4b), (2.4d), or (2.4g) for anyone). As I pointed out in the Introduction, the underlying complexity/acceptability correlation may be obscured through rule-generalization and other phenomena typical of language change, and I believe that the restriction on relative pronoun deletion in non-subject position in certain dialects of Modern English is a case in point. I also expressed the view that the lack of transparency of a posited underlying complexity/acceptability correlation should not discourage efforts aimed at proving the existence of such a correlation; in the case under discussion, supporting evidence for the relevance of (2.1) to (2.3) and (2.4) has in fact been found.

Bever points out that many subjects judge (2.5a) to be just as bad as the starred sentences in (2.4) *on a first presentation*.

(2.5) a. The horse raced past the barn fell.
b. The horse seen past the barn fell.

However, when it is pointed out to them that (2.5a) should be interpreted similarly to (2.5b), they judge (2.5a) as entirely acceptable. It appears therefore that (2.1) is grammatized only when the verb is *unambiguously* tensed. The fact that subjects nevertheless tend to reject (2.5a) on the first presentation strongly suggests that a strategy like (2.1) is operative, for otherwise we would expect (2.5a) to be rejected only about 50 per cent of the time, depending on whether the hearer happened to interpret *raced* as a tensed or an untensed form.[4]

Another experiment cited in Bever (1970) as supporting evidence for (2.1) concerns the two pairs of second degree center-embeddings in (4.6).

(2.6) a. *The editor the authors newspapers* hired liked laughed.
b. *The editor authors the newspapers* hired liked laughed.

Subjects were asked to paraphrase sentences like the above, and experienced much greater difficulty with the b- than with the a-type. This can be again accounted for by (2.1), for, while both construction types involve two interruptions, as well as perceptual conflict, the b-sentence also invites erroneous closure (the underscored string being a possible sentence in (2.6b) but not in (2.6a)).

In an interesting recent paper, Chapin, Smith and Abrahamson (1972) suggest that Bever's strategy is in fact a particular instance of a more general strategy which they formulate as follows:

---

[4] This is, unfortunately, not quite correct, for there are additional variables which viciate this conclusion. Thus, (a) the causative reading of *race*, is somewhat less common than the non-causative reading, and (b) the causative reading requires that the verb be passive, which, as has already been shown experimentally, is a more complex form than the corresponding active one. Assuming a strategy by which, given two options (and all other things being equal), subjects tend to select the less complex option (see also principle (2.22) below in the text), it is conceivable that this strategy may also determine the perception of *raced* (2.5a) as a tensed verb, over and above (2.1); consequently, the initial rejection of (2.5a) cannot be unambiguously attributed to (2.1).

This shows the enormous difficulty of constructing meaningful experiments for syntactic perception, for the huge number of interacting cues makes it extremely hard to control all the pertinent variables.

(2.7) In imposing an initial structural description on a sentence, the S(ubject) attempts at each successive point to close off a constituent of the highest level possible.

They argue that (2.7) can account for all the cases covered by (2.1), and also for the following experimental results which they report on: Subjects were presented with sentences like those in (2.6c) and (2.6d); in the former, the first major boundary is the main break. Clicks were located midway between the two boundaries, and subjects were asked to report the location of the clicks.

(2.6) c. The Chairman of the Board should give whoever invented that process a prize.
d. *Women who are over sixty years old* and kids must have the first chance to escape.

In sentences like (2.6c), there was a significantly greater tendency to prepose the click towards the earlier boundary than to postpose it towards the later one; this was interpreted as a tendency for clicks to migrate towards the main sentence break. On the other hand, in sentences like (2.6d), there was no greater tendency for the clicks to postpose than to prepose, although the main break was the second one; Chapin et al. propose that the factor which interferes with the attraction exerted by the main sentence break is precisely (2.7), which prompts subjects to assume that the underscored string in (2.6d) is an argument of the main clause, while it is in fact only a term of a coordinate argument of the main clause.

Another instance of perceptual complexity explainable by (2.7), but not by (2.1), is (2.6e).

(2.6) e.??*Artists girls, boys and women* hurry a lot tend to resent it.

This sentence was mentioned by Bever in a public lecture at the Linguistic Institute in 1971. This sentence is bad because the string ending with *lot* tends to be misconstrued as the main clause; however, (2.1) cannot tell the whole story here, for the deletion of non-subject relative pronouns is perfectly possible in principle. Thus, *artists girls hurry a lot tend to resent it* is perfectly acceptable, presumably because the lack of pause between *artists* and *girls* prevents a coordinate reading. The cue constituted by the absence of pause after *artists* in (2.6e) is apparently overridden by another strategy, which Bever formulated roughly as "a sequence of the form NP $*\begin{Bmatrix}\text{and}\\ \text{or}\end{Bmatrix}$ NP is interpreted as a coordination." Observe, however, that (2.7) can account for the fact that the underscored sequence in (2.6e) tends to be perceived as a coordinate NP, without requiring an additional strategy: this sequence can be misconstrued as subject of the main clause, and premature closure occurs after *women*, when in fact it should have occurred after *lot*. Thus, as the underscored sequence has been misconstrued as the subject of the main clause, it plus the sequence *hurry a lot* can be misconstrued as the main clause (it appears that (2.7) applies twice in this sentence).

It should be pointed out that (2.7) is inadequate as it stands: If subjects indeed try to close off a constituent of the highest level possible as soon as possible, then why should they close in (2.6e) after *women* and not after *artists*? After all, the latter is a perfectly possible subject of some main clause! Similarly, Chapin et al. argue that the most natural order of prenominal adjectives in English (which, as pointed out in Vendler, 1961, requires that adjectives be arranged from left to right as they increase in "nominality") maximizes the efficiency of (2.7), since it allows the NP to be closed as soon as an element of lesser nominality than the preceding one is encountered. Thus, *the delightful red plastic cup* is more acceptable than ?*the delightful plastic red cup*, because, according to Chapin et al., the latter invites erroneous closure after *plastic*, however, why does closure occur after cup in the former, rather than after *plastic* or even *red*, for both *the delightful red plastic* or *the delightful red* are theoretically possible NP's? It appears that there is something missing in (2.7), and that an amended formulation something like (2.8) would be in order.

(2.8) In imposing an initial structural description on a sentence, the S(ubject) attempts at each successive point to close off a constituent of the highest level possible; however, closure is suspended until some significant cue is encountered.

What can constitute a "significant cue" remains to be investigated for each construction type. In the case of the natural order of adjectives, that cue apparently is the occurrence of a constituent of lesser nominality than the preceding one, which acts as a NP-boundary marker; in coordinations, a helpful cue may be the presence of a coordinating conjunction, which signals that the last member of a coordination is about to come (or has already occurred, in languages like Latin); however, this cue is not available in languages which lack an overt morpheme *and* and conjoin by juxtaposition. In any event, closure must be suspended until *some cue* carrying *sufficient weight* occurs, or coordinations should be unacceptable, optional sentence-final adverbial modifiers should create perceptual problems, and sentences like John screamed at me or *Bill is eating an apple* should receive low acceptability ratings because *John screamed* and *Bill is eating* are possible English sentences.

Before closing this section, it is appropriate to try to pinpoint some factors which may tend to favor the grammatization of complexities resulting from erroneous closure (of course, complete and rigorous predictions are, as I have stressed a number of times, strictly a goal for the remote future). Specifically, we may ask why (2.5a) "improves upon acquaintance," while (2.4b) does not; that is, why the former becomes acceptable once it has been pointed out to the hearer that *raced* is a past participle (as laid down in the rule-of-thumb I gave at the end of section 2.0.1). My guess is that grammatization tends to be inhibited when there exists some overt feature of the string on which the mistaken initial interpretation can be blamed. In (2.5a), the blame can be laid on the hearer's failure to select the correct homonym for the item *raced*. In the starred sentences of (2.4), however, the hearer can only be blamed for his failure to reconstruct an item whose logical presence is not signalled by any overt cue; as the hearer has not misinterpreted any of the overt features of the string, there is no way in which he can devise a strategy for avoiding erroneous closure in the future (i.e., there is nothing to which he can try to pay better attention), and it is reasonable to assume that such sentences will tend to be avoided (and ultimately thrown out of the language).

Selection of the wrong homonym is certainly not the only mistake which may be corrected by paying better attention to some overt cue in the acoustic signal. Thus it is to be expected that suprasegmental cues may also inhibit grammatization; for example, it is possible to distinguish intonationally between [*John or Bill or Mary*] and [*John or* [*Bill or Mary*]] (essentially, by uttering the lower level coordination more quickly). This intonational option presumably enables (2.6e) to improve upon acquaintance, as it should be possible to make it clear by pitch, stress and/or use of pauses that only the last three of the first four NP's form a coordination. On the other hand, it is hard to see what intonational cues could keep the two putative readings of the underscored strings in (2.4) apart: since only nonrestrictive relative clauses allow the deletion of the pronoun in English (compare *the man I hit is sick* with *the man, I hit, is sick) there is no way to signal the intended relative pronoun by pause; on the other hand, short subjects cannot be separated by pause from their verbs either, so that pause cannot distinguish between the two readings here. Notice also that the relation the remainders of the underscored strings bear to the initial NP is exactly the relation they bear to the deleted pronoun, since both elements are subjects; as these remainders would be VP's of the same type on either reading, there is no way in which their suprasegmental structure could force one reading rather than the other. Notice also that the two putative subjects can *always* allow the same VP's (since they are coreferential and therefore necessarily enter into the same selectional restrictions), while erroneous closure in sentences like (2.6e) is only possible when the verb occupying the position of *hurry* can be construed as either transitive or intransitive (compare (2.6e) with the more acceptable *artists girls, boys and women hate the guts of usually develop suicidal tendencies*). I assume that the fact that there are unambiguous sentences constructed on the pattern of (2.6e) except for the lack of an intransitive reading for the counterpart of *hurry* may also inhibit the tendency to throw (2.6e) out of the language.

I have discussed several linguistic illustrations of the operation of a principle of erroneous closure (tentatively formulated as (2.8)), and have also attempted to pinpoint some of the factors that may favor grammatization. Additional (mostly grammatized) instances of erroneous closure will be considered in the ensuing sections and chapters, and I shall argue that their grammatization can be accounted for essentially by the same principle I have postulated here, namely, that erroneous closure tends to result in ungrammaticality when it cannot be avoided by concentrating on some overt feature of the string (which may have been overlooked on the first take).

2.1.2. This section is devoted to the so-called "cross-over phenomena" discussed in Postal (1971). Postal recognizes two broad subclasses of phenomena, namely, cross-over restrictions on bounded rules, and on unbounded rules. I shall argue that the former subclass is essentially explainable in terms of a conflict principle--(2.23)--, and that most of the phenomena in the latter can be explained on the basis of an _interruption principle_--(2.37)--coupled with the already mentioned principle (2.22), which, in strategic terms, translates as the instruction "whenever possible, choose the simplest alternative."

In his book, Postal presents the cross-over constraints as purely syntactic, on the grounds that the sentences they rule out "have unique interpretations and involve no semantic violations at all." Obviously, Postal did not consider the possibility that his phenomena might have a psychological explanation, but I shall argue that such an explanation can be both simpler and more illuminating.

The bounded rules that Postal is concerned with are Passive, Subject-Raising, Tough-Movement, About-Movement and Flip, and these are illustrated in (2.9)-(2.13) respectively.

(2.9) a. _John_ stabbed _Mary_.
b. _Mary_ was stabbed by _John_.

(2.10) a. It seems to _me_ that _John_ is stupid.
b. _John_ seems to _me_ to be stupid.

(2.11) a. It is difficult for _Mary_ to shave _Bill_.
b. _Bill_ is difficult for _Mary_ to shave.

(2.12) a. I talked to _John_ about _Mary_.
b. I talked about _Mary_ to _John_.

(2.13) a. _I_ am amazed at _John_.
b. _John_ amazes _me_.

Postal points out that in his idiolect--and in a large number of dialects--the b-sentences in the above set become ungrammatical if the underscored NP's are coreferential, as in (2.14)-(2.18b).

(2.14) a. _John_ stabbed _himself_.
b. *_John_ was stabbed by _himself_.

(2.15) a. It seems to _me_ that _I_ am stupid.
b. *_I_ seem to _myself_ to be stupid.

(2.16) a. It is difficult for _me_ to shave _myself_.
b. *_I_ am difficult for _myself_ to shave.

(2.17) a. I talked to _Bill_ about _himself_.
b. *I talked about _Bill_ to _himself_.

(2.18) a. I am amazed at myself.
b. ?I amaze myself.

At first blush, we seem to be dealing with a purely syntactic restriction, which prevents bounded movement rules from causing an NP to cross over another NP with which it is coreferential. However, there are several problems with a formal statement along these lines.

First, the existence of some of the transformations claimed to be involved in the above set has been thrown into doubt by various writers. Thus, Jackendoff (1969), offered some evidence against Tough-Movement, and Postal himself points out that he does not claim that (2.18b) is derived from (2.18a), or that sentences related by Flip are necessarily synonymous. Other writers have pointed out that if (2.13a) and (2.13b) are claimed to be related transformationally, and if a is claimed to be in some sense more basic than b because its surface subject is its logical subject (an experiencer), there are other pairs of predicates which, like buy and sell, express converse relations, and still there are no obvious reasons for choosing one of them as more basic than the other. In addition to this, the existence of a restriction in Postal's Flip-cases is not nearly as clearly established as for the other four cases; indeed, most people that I have consulted find (2.18b) perfectly acceptable, and this is also true of other sentences that Postal stars, e.g., I am acceptable to myself or I am loathsome to myself. Still, my informants found other examples of Postal's less acceptable, e.g., I strike myself as pompous. Be this as it may, people usually find (2.17b) just as bad as (2.14)-(2.16b), and if it should turn out that there is no Tough-Movement rule, the entire cross-over principle would be thrown into doubt.

Second, Postal points out an interesting exception to the restrictions on About-Movement: the restrictions do not hold if the sentence contains an earlier NP coreferential with the two NP's that cross over each other, as in (2.19).

(2.19) a. I talked to myself about myself.
b. I talked about myself to myself.

In order to handle this case, Postal invents the notion of pronominal virgin, i.e., an NP which has not yet been pronominalized at the stage at which some transformation applies, and requires that the cross-over restrictions apply only to pronominal virgins. Then, the desired result follows, if Reflexivization is ordered before about-movement. Postal uses the notion of pronominal virgin also in attempting to account for restrictions on unbounded movements, which do not affect reflexives, but do affect simple pronouns; his solution is to order Pronominalization after about-movement. This proposal assumes that Pronominalization is a rule (which has been seriously questioned lately), and makes a number of assumptions about rule ordering which the author himself later regarded as "doubtful" (in Postal, 1970). Moreover, this solution gives no insight into the possible reasons for the constraints' behaving as they do.

Third, Postal notes that the constraints do not hold if the second NP is heavily stressed, and especially "if such particles as even are added" (p. 9). Thus, compare the sentences in (2.20).

(2.20) a. *John is hated by himself.
b. John is hated by himsélf.
c. John is even hated by himsélf.

Postal points out that his formulation of Cross-Over does not account for the acceptability of (2.20b, c) and devotes a separate chapter (Chapter 19) to a discussion of this problem. Basically, he proposes that coreference can either be binding or predicational, and that the latter kind (illustrated by (2.20b, c)) is exempted from the constraints. Binding coreference arises when the speaker presupposes two NP's to be coreferents, but the sentence is about something else, while predicational coreference arises when no identity is presupposed and when asserting coreferentiality between two NP's is the comment of the sentence. While this account may well be correct, it is not obvious why Cross-Over should single out binding coreferentiality as its domain of application.

We shall now consider a possible explanation for the unacceptability of the b-sentences in (2.14)-(2.18). Let us begin by noticing that the b-sentences in (2.9)-(2.13) are psychologically more complex than the corresponding a-sentences (or, to use Prague-school terminology, the a-sentences are "unmarked," while the b-sentences are "marked"). Indeed, in the case of the passive, it is clear that it is more complex than the active, since children acquire it much later and also, many adults tend to avoid passives, whenever possible). Bever suggests that the greater complexity of the passive is due to the violation of (2.2) (in fact, Bever's "actor-action-object" should be replaced by something like "logical subject-predicator-logical object," to cover cases in which the subject is inanimate, and therefore an instrument, or in which the verb is stative and the subject an experiencer). I do not know what the relative order of acquisition is for the members of the pairs (2.10)-(2.13), but if the complexity of the passive is indeed due to a violation of (2.2), then a similar violation occurs in (2.13b). Indeed, only the NP-slot filled by me is necessarily animate, as can be seen in (2.21).

(2.21) a. I am amazed at the scenery.
b. *The scenery is amazed at John.
c. The scenery amazes me.
d. *John amazes the scenery.

With respect to the b-sentences in (2.10) and (2.11), they are more complex than the a-sentences for they involve logical discontinuities in surface structure. Indeed, the predicate seem takes two arguments, an experiencer and a sentential object. The latter forms a surface unit in the a-sentence (that John is stupid), but is broken down into two separate parts in the b-sentence (John and to be stupid). Similarly, the predicate difficult also takes an experiencer and a sentential object (I assume, following Jackendoff, 1969, and Chomsky, 1971, that an earlier stage in the derivation of (2.11a) is it is difficult for Mary$_i$ for Mary$_i$ to shave Bill). Like the sentences in (2.10), (2.11a) exhibits a continuous sentential object, *(for Mary) to shave Bill*, while (2.11b) exhibits a discontinuous one, *Bill* and *(for Mary) to shave*. Finally, we can argue that (2.12) is unmarked because the linear order of its three NP's expresses the natural hierarchy of their roles, since *John* is an experiencer, and therefore necessarily animate, while *Mary* is an object and can be replaced by inanimate or abstract nouns, as well as by nominalizations; on the assumption that animate roles are higher in the hierarchy than inanimate ones, the failure of the linear order of (2.12b) to reflect this fact makes it more marked than (2.12a).[4]

---

[4] It would be tempting to generalize from these facts and claim that the unmarked order of arguments *always* reflects the hierarchy of their roles. However, there are obvious difficulties with such a proposal. For example, the unmarked ordering of Instrumental and Objective differs according to construction, as in (i):

(i) The man broke the window with a hammer.
(ii) *The man broke with a hammer the window.
(iii) The manner broke the window.
(iv)??The window broke the hammer.

As we can see, (ii) is ill-formed, and (iv), even though it is probably well-formed grammatically and odd for non-linguistic reasons, is not a paraphrase of (iii).

Furthermore, it is not even clear whether animate cases (i.e., Agent and Experiencer) necessarily precede all other cases in unmarked order. Thus, it is not clear that (v) is less complex than (vi).

(v) I gave John a book.
(vi) I gave a book to John.

Perhaps, then, the feeling that (2.12a) is less complex than (2.13b) is merely due to the former's greater frequency of occurrence in English. In that case, the phenomenon would be no more remarkable than that place adverbs usually precede time adverbs in English, while the opposite situation obtains in other languages, like French or Rumanian. Be this as it may, the explanation of cross-over phenomena that I wish to propose depends only on the *fact* of markedness and not on its *cause*. Notice, incidentally, that the more general strategy (2.2) must also be language-specific, as one would not expect it to develop in a language with unmarked surface order VOS [apparently, Malagasy is such a language, according to E. Keenan (public lecture, 1971 Linguistic Institute)].

Now that the facts of markedness are established for the pairs in (2.6)-(2.13), it is important to notice that if the members of the various pairs are *cognitively synonymous*, they differ with respect to *focus*.[5]

---

[5] The term 'focus' is often used in the literature as a synonym of 'comment,' and is thus contrasted with 'topic.' However, I use 'focus' in this section to mean 'that which is more prominent subjectively.' Thus, in (i) below, the focus coincides with the topic (*John*); however, in (ii), the focus coincides with the comment (*a snáke*, the stress being contrastive).

(i) John saw a snake.
(ii) John saw a snáke.

This use of the term 'focus' is also found in Partee (1971), who notes that 'adding [contrastive] stress [an instance of comment] and raising to subject [an instance of topic] are two ways of accomplishing the same result, namely the *bringing into focus*' (underscore mine).

As I have defined it, focus is a relative notion, and there may be more than one focus per sentence. In fact, the explanation of the grammaticality of (2.19b) which I propose below rests precisely on the assumption that sentences may be plurifocal.

Change of focus can be achieved in a variety of ways (heavy stress, clefting, unbounded topicalization rules, topic-introducers like *as for*, *with respect to*, etc.), and in the cases at issue, it can be achieved through the use of a marked form, which causes an *earlier* appearance of some argument with respect to the unmarked form, thereby throwing this argument into focus and defocusing the argument which constitutes a focus in the corresponding unmarked form. Thus, Passive and Flip focus on the logical object and defocus the logical subject, Subject-Raising and Tough-Movement focus an element of the sentential object and defocus the experiencer, and About-Movement focuses the *about*-phrase and defocuses the experiencer. Clearly, however, as focus is created by displacing arguments *within a sentence*, it is only the *earliest* appearnace of some NP in that sentence which can have focusing effect (the pertinence of this fact will become obvious below, when (2.19b) is discussed).

I shall now introduce two fairly obvious principles based on experience, which, I believe, can explain not only the problems connected with the bounded rules, but also with the unbounded ones.

(2.22) Unless there is an indication to the contrary, one tends to assume that people do not complicate matters unnecessarily. [This principle is a reformulation of a conversational implicature proposed in Grice (1968)].

(2.23) When something is done for a purpose, and that purpose cannot be achieved in principle, perceptual conflict arises.

(2.23) can provide a simple explanation for the unacceptability of the b-sentences in (2.14)-(2.18). Indeed, a hearer naturally assumes that a speaker uses the more complex marked form in order to focus on some NP and defocus another; however, there turns out to be no such possibility in these cases, since the focused and defocused NP's refer to the same entity. In other words, the use of the marked form is pointless.

To put the matter another way, the sentences in question are odd because they create conflicting assumptions in the hearer; the fact that the marked form is used leads the hearer to believe that the focused and defocused entities are different; on the other hand, the existence of a reflexive pronoun tells the hearer that the speaker presupposes that the two entities are not different. The reason why these sentences are odd but interpretable, is that the assumption of coreferentiality alone is a logical presupposition, while the assumption of non-coreferentiality is merely an implicature. If both assumptions were logical presuppositions, the sentences would be semantically ill-formed, like my brother, Mary, is sick.

Let us take a quick look at the various problems that I earlier claimed arise if the cross-over phenomena are treated in purely formal terms.

First, the cross-over principle, as formulated by Postal, depends crucially on there existing a movement transformation. Nothing of the kind is required for a principle like (2.23) to be applicable. Indeed, all we need is the existence of pairs of cognitively synonymous sentences, which differ in perceptual complexity and in focus potentiality; the synonymy need not, however, be statable transformationally. I believe that the sentences related by the putative Tough-movement rule are synonymous (at least in the vast majority of cases, since the counterexamples advanced in Jackendoff (1969) are sufficiently rare for us to be able to assume that some learners may never hear them). Non-synonymy, on the other hand, seems to be the rule rather than the exception for 'flipped' sentences, and this is probably why the constraints are very much weaker or totally non-existent for most speakers.

Second, in order to account for the acceptability of (2.19b), Postal had to invent the notion of pronominal virgins and stipulate that the cross-over restrictions apply to pronominal virgins only; this in turn called for a number of assumptions about rule-ordering. Notice, however, that (2.19b) is not ruled out--or marked as odd--by either (2.23) or (2.22). Indeed, due to the fact that there is an earlier NP--the subject--which is coreferential with both reflexive pronouns, the about-phrase fronting can no longer have a focusing effect, and therefore there is no assumption that it took place 'for a purpose,' as required for the applicability of (2.23). The choice of the marked form is still an 'unnecessary complication' (see (2.22)), and one might assume that the two NP's which switched places are not coreferential, if it were not for their reflexive form, which is 'an indication to the contrary.' Thus, (2.19b) is possible, as a stylistic variant of (2.19a).

Third, Postal attempts to explain the failure of heavily stressed reflexives to obey the constraints by claiming that the NP's involved are not presupposed to be coreferential at the stage at which their coreferentiality is asserted. Specifically, he derives (2.24a) from (2.24b) by a rather complex transformation which he calls Contrast-Movement.

(2.24) a. Bill cut himsélf.
b. The one$_j$ [who$_j$ Bill$_i$ cut] was Bill$_i$ (himself).

I believe that if (2.24a) and (2.24b) are related transformationally, it is the former, rather than the latter, which is basic, and this claim will be defended in considerable detail in 2.1.3. The directionality of the transformation relating (2.24a) and (2.24b) is however not crucial here, for in either case *Bill* and *himself* are not presupposed to be coreferential. One possible explanation of the acceptability of *John was shot by himsélf* is that the focus of a sentence is in most cases its subject, therefore a change of subject by one of the bounded rules at issue effects a change of focus; however, if there is a stressed constituent in the sentence other than the subject, *that one* counts as the focus. Since the subject is no longer the focus, it can no longer be assumed that the change of subject was made in order to create a new focus; rather, it is assumed that the marked form was selected for some other reason(s). The reason which determines such selection in a case like *John was shot by himsélf* is that this sentence presupposes *John$_i$ was shot by someone$_j$* (usually due to a previous assertion, such as *John was shot by Bill*); thus, *John wasn't shot by Bill, he shot himsélf* sounds much less natural than *John wasn't shot by Bill, he was shot by himsélf*. Notice that this line of approach enables us to explain why **Jóhn was shot by himself* is bad; indeed, this sentence presupposes **someone$_i$ was shot by himself$_i$*.

We are also in a position to understand why (2.20c) is better than (2.20b), a fact which Postal's Contrast Movement rule cannot account for. Indeed, the scope of *even* constitutes a focus, and (2.20c) has two overt signals that the subject is not the focus, namely, *even* and contrastive stress, while (2.20b) has only one such signal, contrastive stress. If the crossover phenomena are due to behavioral reasons, as I suggested, then it is quite natural to expect that two signals should be better than one in this situation, for they make the shift in focus easier to perceive.

In stating Cross-Over, Postal restricts its applicability by using the notions of mention, peer, clause-mate (we have already mentioned pronominal virgin). While these notions do indeed yield the desired results, it remains to be shown whether there is a deeper explanation why precisely these notions should have to be used in precisely the ways in which Postal uses them. In other words, it is legitimate to ask whether these notions are in some way connected with the fact that there are cross-over constraints, or whether they constitute a set of arbitrary, unrelated consitions. I believe that the notions in question are natural consequences of the account I have proposed.

The requirement that only NP's mentioned in the structural description of a transformation be subject to cross-over restrictions (originally made in Ross, 1967), is a totally superfluous requirement in the framework I am proposing. Consider for example (2.25b), which is acceptable--according to Postal--because *him*$_i$ is not mentioned in the structural description of Subject-Raising, which moves it over *John*$_i$. Observe, however, that (2.23) does not mark (2.25b) as odd, because the focus of the unmarked form is *John*, while the focus of the marked form is *the woman who loves him*, and the attempt to create a new focus does not fail.

(2.25) a. It seems to John$_i$ that the woman who loves him$_i$ is sick.
b. The woman who loves him$_i$ seems to John$_i$ to be sick.

Similar arguments will be seen to apply in the case of the unbounded rules.

The notion of peer is meant to account for the acceptability of (2.25b), where the two coreferents are not peers.

(2.26) a. Johni's wife struck him$_i$.
b. John$_i$ was struck by his$_i$ wife.

This specification is again unnecessary in my account, for the unmarked form focus is *John's wife*, while the marked form focus is *John*, and (2.23) is inapplicable.

Finally, Postal's Clause-Mate Restriction must apply to NP's which are clause-mates *before* one of them crosses over the other, in order that the crossing movement may be blocked. This causes serious difficulties in relation to Subject-Raising (or It-Replacement, as Rosenbaum called it), because the coreferents are not clause-mates in the input to this rule. Postal proposes to obviate this difficulty by breaking It-Replacement into two steps, one which raises the complement subject to object position in the matrix sentence and one which flips the new object to subject position. Thus, rather than derive (2.15b) from (2.15a), he proposes that the source of these sentences is *to me seems that I am stupid*, which yields *to me seems I to be stupid*, and ultimately **I seem to myself to be stupid*. This account fails at the level of observational adequacy, since speakers do not, in general, have the same intuitions about the severity of violations arising through Subject-Raising and through clear cases of the putative Flip rule.

In my account, if a clause-mate requirement were needed, it would be needed only in the *output* of the various putative transformations, a condition fulfilled in all of (2.14)-(2.18). This is so because the applicability of (2.23) can only be judged with respect to the marked form. For (2.23) to reject a marked form, the new focus must of course be a clause mate of the unmarked form focus, since focus by bounded rules is precisely created *by causing an NP to occur as an earlier argument of some verb*.

We shall now consider the restrictions on unbounded rules, which I shall try to show are explainable, to a large extent, by (2.22). Let us briefly consider the facts and Postal's solution.

Postal concentrates on three unbounded rules: Wh-Q-Movement, Wh-Rel-Movement, and Y-Movement. I shall also consider Clefting, which is interesting to compare with Y-Movement, on account of formal similarities. The data to be explained are given in the paradigm below.

(2.27) a. Mary$_i$ thinks that Bill loves her$_i$.
b. Who$_i$ thinks that Bill loves her$_i$?
c. *Who$_i$ does Mary$_i$ think that Bill loves?

(2.28) a. The girl$_i$ [the girl$_i$ knows Bill loves the girl$_i$] is pretty.
b. The girl$_i$ who$_i$ knows Bill loves her$_i$ is pretty.
c. *The girl$_i$ who$_i$ she$_i$ knows Bill loves is pretty.

(2.29) a. Mary emphasized John$_i$ must think that Jill loves him$_i$.
b. John$_i$, Mary emphasized must think that Jill loves him$_i$.
c. *Him$_i$, Mary emphasized John$_i$ must think that Jill loves.

(2.30) a. John talked to Mary about herself.
b. Mary, John talked to about herself.
c. Herself, John talked to Mary about.

(2.31) a. John$_i$ thinks that Mary loves him$_i$.
b. It's Jóhn$_i$ who thinks that Mary loves him$_i$.
c. It's hím$_i$ John$_i$ thinks Mary loves.

(2.32) a. John talked to Mary about herself.
b. It's Máry John talked to about herself.
c. It's hersélf John talked to Mary about.

(2.33) a. *[Whose$_i$ mother] does John$_i$ admire?
b. *The man$_i$ [whose$_i$ mother] he$_i$ hates is a freak.
c. [His$_i$ mother], John$_i$ admires.

One similarity between the restrictions on bounded and unbounded rules is that neither seems to be universal. This is probably a consequence of the fact that both (2.22) and (2.23) are *perceptual principles*, but not *logical necessities*.

Another similarity between the two kinds of rules is that restrictions apply only to NP's mentioned in the index of those rules. Thus, Postal points out Wh-Q-Movement is free to operate in (2.34), because it is the *wh*-word, and not *him*, which is mentioned in the index of these transformations.

(2.34) a. Charley$_i$ visited some of the men who criticized him$_i$.
b. [Which of the men who criticized him$_i$] did Charley$_i$ visit.

In contrast to bounded rules, unbounded rules obey the constraints even when the crossing does not involve clause-mates, either in the input or in the output, as can be seen in (2.27)-(2.29).

Another difference is that crossing restrictions are not always restricted to peers. As can be seen in (2.33), the constraints hold even for non-peers in the case of Wh-Q-Movement and Wh-Rel-Movement, although not in the case of Y-Movement. In other words, even though the moved NP_i is not a coreferent of any other NP_j that it crosses over, unacceptability may arise if the moved NP_i contains an NP_k coreferential with NP_j. However, why should unacceptability arise in (2.33a) and (2.33b) but not in (2.33c)? Postal invokes Ross' notion of Pied-Piping, which says that if a wh-word is part of a larger NP (and if some additional conditions are satisfied--for which, see Chapter One, (1.22)--) the larger NP *may* be moved along with the wh-word; if, however, the wh-word is on a left branch, the larger NP *must* accompany the wh-word. As (2.33a) and (2.33b) have the wh-word on a left branch, Postal proposes to regard the *wh*-word as the mentioned NP, and to ascribe the fact that it did not move alone to the LBC. On the other hand, Postal claims that Pied-Piping is not applicable to Y-Movement, and thus (2.33c) is grammatical because *his* cannot be the mentioned NP. However, Postal does not ask why Pied-Piping should be inapplicable to Y-Movement. Notice that Pied-Piping does not depend on a wh-word; rather, it seems to require that the mentioned NP should be identifiable by some overt surface mark, thus suggesting that perception is somehow involved. This hypothesis is supported by the fact that Pied-Piping *is* applicable to Clefting, which differs from Y-Movement only in allowing the identification of the clefted element by contrastive stress. Thus, it is reasonable to regard *his mother* as the mentioned NP in (2.33d), but *his* as the mentioned one in (2.33e).

(2.33) d. It's [his_i móther] $John_i$ admires most.
e. It's [$hís_i$ mother] $John_i$ admires most.

A third difference between the two types of rules is that none of the starred sentences in which cross-over violations by unbounded rules occur is unacceptable *on all its readings*. Thus, for every starred sentence in (2.27)-(2.33) there is a reading on which the crossed-over NP's bear different indices, i.e., they are not presupposed to be coreferents (which does not mean that they are presupposed to be non-coreferential, for a perfectly natural answer to (2.27c) is *Mary herself*; also, the following is semantically well-formed: *I know who Mary thinks Bill loves, it's Mary herself*, etc.). Moreover, when the moved NP is a reflexive as in (2.30) the coreferential interpretation is both semantically possible and acceptable (or grammatical); Postal accounts for this fact by pointing out that reflexivization precedes unbounded movements, and that the moved reflexive is not a pronominal virgin.

Finally, it should be pointed out that there are no cross-over constraints on Clefting--as seen in (2.31)-(2.32)--at least for the dialect in question.

In proposing an alternative to Postal's account, we shall have to answer at least the following questions:

(2.35) a. Why do Wh-Q-Movement, Wh-Rel-Movement, and Y-Movement obey the constraints, but not Clefting?
b. Why do unbounded movements not result in unacceptabilities, and why are reflexives exempt from the constraints?
c. Why are the constraints confined to "mentioned" NP's?
d. Why is Pied-Piping inapplicable to Y-Movement?

With respect to (2.35a), we shall first inquire whether the constraints can be explained by either (2.22) or (2.23). A quick comparison of the paradigm in (2.27)-(2.33) with the one in (2.14)-(2.18) will show us that the choice is between the input and the output of some transformation in the latter, but between two outputs in the former. Out of the four unbounded rules considered, only Wh-Rel-Movement and Y-Movement qualify for either (2.22) or (2.23), since it is only for these rules that the two possible outputs (i.e., the one in which the closest possible candidate has moved and the one in which the more remote has done so) are cognitively synonymous. Indeed, there is every reason to assume that both outputs are derived from the same underlying representation: the Rel-Wh-feature can be placed on any coreferent of the head, without semantic effect, and Y-Movement, which fronts the topic of the sentence, achieves the same effect by moving any of several coreferents. On the other hand, when different coreferents are questioned or clefted, the results are not cognitively synonymous, as the various outputs have different underlying representations and different presuppositions. Indeed, the source of (2.27b) is *someone*$_i$ *thinks that Bill loves her*$_i$ while the source of (2.27c) is **Mary*$_i$ *thinks that Bill loves someone*$_i$. With respect to clefted NP's, they are semantically equivalent to contrastively stressed ones, so that (2.31b) and (2.31c) are related to the semantically distinct *Jóhn thinks that Mary loves him* and *John thinks that Mary loves hím*, respectively. Some confusion may arise between clefted and topicalized sentences, because informants may involuntarily stress NP's fronted by Y-Movement, which makes them indeed equivalent to clefted ones. However, if one is careful not to put constrastive stress on NP's topicalized by Y-Movement, the different behavior of the two rules becomes clear. The paradigm in (2.36) shows that clefted reflexives behave like stressed ones and topicalized reflexives behave like unstressed ones with respect to the constraints.

(2.36) a. *John was cut by himself.
b. *Himself, John was cut by.
c. John was cut by himsélf.
d. It's himsélf John was cut by.

As Clefting does not obey the constraints, there is nothing to explain, and the inapplicability of (2.22) or (2.23) follows from the fact that the various outputs are not cognitively synonymous. The problem will be to show why the constraints hold for questions, since the various outputs are not derived from the same deep structure, or from synonymous deep structures. Thus, questions will have to be explained by a different principle than relatives, which I do not regard as a weakness of my analysis, since there is no logical necessity that questions and relatives receive the same explanation, although Postal assumes such a necessity without comment or justification.

We have established that when several outputs are possible for Wh-Rel-Movement and Y-Movement, these outputs are expected to be synonymous, as they are derived from the same underlying representation. I do not see how the fact that the more remote coreferent cannot be relativized or topicalized can be accounted for by (2.23), since I do not see what could qualify as a 'purpose'. In the case of the bounded rules, the purpose was interpreted as the creation of a new focus, and this is inapplicable here. I will therefore assume that (2.23) is inapplicable, which already provides a partial answer to (2.35b), since (2.23), but not (2.22), can mark a sentence as unacceptable. To see whether (2.22) is applicable in these cases, we must first establish that the choice of a more remote coreferent would 'complicate matters' in some way. I believe that such a choice would indeed constitute a complication, especially in view of the fact that the linear distance and the number of intervening sentence nodes between the two candidate NP's could be arbitrarily large, since the rules involved are unbounded ones. Indeed, the choice of the more remote constituent would unnecessarily strain the hearer's memory, by forcing him to await the 'resolution' of the fronted constituent--i.e., its 'plugging' into the constituent to which it immediately belongs, thereby revealing one of its roles[7]-- for a longer period of time. I do not know whether the greater complexity involved in the choice of a more remote coreferent is reflected in the process of acquisition; that is, I do not know whether children first learn to relativize or topicalize out of the first sentence down, then out of the second, and so on. Such complexity has, however, experimentally been shown to exist for the adult due to the following interruption principle (Bever, 1970);

(2.37) Discontinuous components are perceptually complex in proportion to the structural complexity of the intervening material.

Notice that in the output of neither rule is there an indication that the two coreferents must be coreferential, as personal pronouns need not be coreferential with anything but the relative clause head. Therefore, the explanation of the unacceptability of (2.27c) and (2.29c) is quite simple: given the fact that discontinuities increase in complexity with the complexity of intervening material, there is a reasonable assumption that discontinuities will be as 'small' as possible; as there is no 'indication to the contrary', i.e., no overt mark of necessary coreferentiality, a hearer assumes that the discontinuities in the two above-mentioned sentences are minimal, and that there is no coreferent of the fronted NP between its surface and underlying positions. The situation is different with respect to (2.30c), as the fronted reflexive does constitute an overt mark of necessary coreferentiality. The hearer knows that a fronted constituent plays two roles (as pointed out in footnote 7). He also knows that a reflexive must have a clause-mate coreferent. As no such clause-mate can exist in its surface position, the conclusion is unavoidable that it must exist in its underlying position. As the hearer also knows that the clause-mate must be to the left of the reflexive, a topicalized reflexive already tells him that the latter must have 'crossed' a coreferent. Therefore, the coreferential interpretation is possible (in fact, necessary) because (2.22) is inapplicable, and this interpretation is well-formed because (2.23) is inapplicable. Notice that the explanation we have proposed has also provided an answer to (2.35b).

---

[7] Fronted constituents play at least two roles: they serve as complementizers or topic-introducers in their surface position, and as arguments of the predicate of some lower clause in their underlying position.

Principle (2.22) can explain an interesting dialectal peculiarity. Thus, the paradigm in (2.38) holds for some speakers:

(2.38) a. Mary told John_i that she loves him_i.
b. *John_i, Mary told that she loves him_i.
c. *Him_i, Mary told John_i that she loves.
d. John, Mary told that she loves.

Neither of the two coreferents can move *alone* for these speakers, under conditions that I have not investigated in detail. For our purposes, it is sufficient to notice that chopping must occur 'across-the-board' when both coreferents occupy object position. Ross (1967) pointed out that the across-the-board condition constitutes an absolute exception on unbounded chopping rules (in the sense of Lakoff, 1970a) in coordinations. That is, it is not only the case that chopping must occur in all the terms of a coordination *if* each contains a coreferent, but also each coordinate term *must* contain a coreferent in specific positions if chopping is to occur at all. We can see that the across-the-board condition overrides the strategy which requires that unbounded movements be as economical as possible.

While the across-the-board condition in coordinations is a very strong, and possibly universal, one, there appears to be a weaker condition with respect to other structures; moreover, the strength of this condition varies dialectally. Thus, for the dialect in question, across-the-board chopping is mandatory in structures like (2.38a), but *only if* the matrix and the complement contain coreferents. Thus, we are not dealing here with an absolute exception, for *John, Mary told that she loves Jack* is perfectly well-formed; in contrast, **John, Mary loves and Jean hates Bill* is not. Moreover, (2.38b) is not ill-formed for the speakers in question on all its readings, but only with the linkages indicated. It is interesting that if the linkages indicated are impossible in (2.38b) in the dialect at issue, there are many speakers who accept (2.38b), but find it odd and prefer (2.38d) instead. Ross (1967) mentions a number of adverbial constructions which differ from both cases mentioned above in that chopping is precluded out of the adverbial, but possible out of the matrix or out of both; however, if the environment for across-the-board chopping exists, chopping out of the matrix alone is sometimes odd; more specifically, the acceptability of chopping out of the matrix alone when the adverbial contains a coreferent seems to be inversely proportional to the acceptability of across-the-board chopping in the same cases.

We shall return to across-the-board chopping later, and will try to provide an explanation for the fact that across-the-board chopping is sometimes an absolute exception and sometimes not, as well as for the fact that it is possible in matrix and adverbial constructions in which the adverbial alone cannot lose an element through chopping under any circumstances. For the time being, we shall only notice that the reasons for the unacceptability (or marginality) of (2.38b) are fairly straightforward: while (2.38d) creates an additional discontinuity, as compared with (2.38b), it also makes it perfectly unambiguous that the object of the matrix and that of the complement are coreferents; on the other hand, (2.38b) leaves the referent of *him* ambiguous, for although the linkages indicated are a possibility in some dialects, they are nowhere a necessity. Thus, (2.38b) appears odd, because the hearer may assume that the speaker will not allow his discourse to be more ambiguous than necessary; in fact, we seem to be dealing with a clash of two strategies: one that assumes that discontinuities are not increased without reason, and one that assumes that ambiguities are not proliferated without reason (both are instantiations of (2.22)). The dialectal variation in the acceptability of (2.38b) reflects the reasonable assumption that neither strategy is overwhelmingly stronger than the other; the way in which the conflict at issue is resolved determines the particular degree of acceptability of such sentences in individual dialects.

The across-the-board condition was illustrated with Y-Movement only in (2.38), but it holds for all unbounded movements. For example, it holds for Clefting (which is in general[8] free from cross-over constraints), as shown in (2.39) (as in the case of Y-Movement, the acceptability of (2.39b) is subject to dialectal variation).

(2.39) a. It's John Mary told that she loves.
b. *It's John$_i$ Mary told that she loves him$_i$.

---

[8] I have found that the speakers who accept (2.31b) reject (2.31c), instead of which they say *It's himself$_i$ John$_i$ thinks Mary loves*. This fact is again explainable by (2.22), since, so long as the reflexive is possible in this situation, it is clearly a less ambiguous, and therefore simpler, choice than the personal pronoun; thus, the use of a personal pronoun creates presumption of noncoreferentiality in this dialect.

The reason is that (2.39a) has the same meaning in the dialects in question as (2.39b) purports to have (or that it has, in the dialects in which it is acceptable). As a choice between two synonymous forms exists, (2.22) forces the choice of (2.39a).

I shall now attempt to provide an explanation for the behavior of questions.

Postal (1971) notes in Chapter 21 that Jackendoff has proposed, at the 2nd Annual La Jolla Conference on English Syntax, that wh-words are *wh+some* words, and that the ill-formedness of (2.40a) is due to the ill-formedness of its source, (2.40b).

(2.40) a. *Who$_i$ did she$_i$ claim Jack kissed?
b. *She$_i$ claimed Jack kissed someone$_i$.

Postal rejects Jackendoff's proposal, but since he later admits in Postal (1970a)[9] that he had done so for 'largely wrong reasons,' I shall not try to show that Postal's initial criticism was mistaken. It will be interesting, however, to look at what Postal regards (in the more recent paper) as 'right reasons'. Postal gives four arguments:

(2.41) a. In order to make the argument go through, it is necessary to show that restrictive and appositive relative clauses obey the constraints *for the same reasons* as questions.
b. It is not true, as Jackendoff claims, that *some*-words can never refer backwards to a personal pronoun, as (2.42a) shows. Moreover, there are a number of cases in which the wh-word is not fronted, e.g., (2.42b) and in those cases the wh-word *can* refer to a preceding pronoun.

(2.42) a. Although he$_i$ thought he$_i$ should run, I explained to some man$_i$ (I know) that the FBI would catch him$_i$ in no time.
b. The newsman who criticized him$_i$ later belted which official$_i$?

(2.41) c. The constraints hold not only for *wh*-words, but also for a certain class of nominals which contain them. But although a sentence like (2.43a), is ill-formed, its source, (2.43b), is well-formed.

(2.43) a. *Whose$_i$ father$_j$'s brother$_k$ did his $\left\{\begin{matrix} i \\ j \\ k \end{matrix}\right\}$ sister criticize?
b. His $\left\{\begin{matrix} i \\ j \\ k \end{matrix}\right\}$ sister criticized the brother$_k$ of the father$_j$ of someone$_i$.

---

[9] Postal (1971) first appeared in 1968.

(2.41) d. In question clauses with more than one *wh*-word, the *wh*-word which does not move can refer backwards to a pronoun. However, if that *wh*-word is fronted, it can no longer refer forwards to that pronoun, as shown in (2.44).

(2.44) a. What company had his$_i$ wife spy on what well-known industrialist$_i$?
b. *What well-known industrialist$_i$ did the company have his$_i$ wife spy on?

With respect to (2.41a), I think it is simply wrong. Indeed, if the same constraints obtain in two separate instances, it might be desirable, but certainly not *necessary*, to show that the same cause determines the restrictions on both cases.

With respect to (2.41b), I think that Jackendoff's claim is too strong, due to his failure to distinguish between the properties of stressed and unstressed *some*. Although I have not conducted an exhaustive investigation of *some*, I wish to propose the following tentative generalization, based on Postal's examples:

(2.45) a. unstressed *some* cannot refer backwards to a pronoun.
b. stressed *some* can refer backwards to a pronoun, subject to the general constraints on backwards pronominalization (i.e., *some* cannot be both commanded and preceded by the pronoun). By way of illustration, consider (2.46) and (2.47).

(2.46) a. That he$_i$ was sick disturbed John$_i$.
b. That he$_i$ was sick disturbed sóme boy$_i$.
c. *That he$_i$ was sick disturbed sŏme boy$_i$.

(2.47) a. That his$_i$ mother was dying disturbed John$_i$.
b. That his$_i$ mother was dying disturbed sómeone$_i$.
c. *That his$_i$ mother was dying disturbed sŏmeone$_i$.

If we assume that question wh-words come from the *some* which ultimately receives stress, rather than from the one which does not, and if Jackendoff's proposal is modified to incorporate this assumption, Postal's argument loses its strength. Indeed, consider Postal's example (36), which I reproduce below as (2.48a) and (2.48b) and which purports to show that questions can be well-formed when the underlying structure that Jackendoff assigns to them is ill-formed; however, if the underlying structure is taken to be (2.48c), rather than (2.48a), we see that the question and its source agree in grammaticality.

(2.48) a. *The tyrant tortured her$_i$ mother in front of sŏme helpless princess$_i$.
b. What tyrant tortured her$_i$ mother in front of what helpless princess$_i$?
c. The tyrant tortured her$_i$ mother in front of sóme helpless princess$_i$.

There are independent reasons for believing that interrogative *wh*-words come from the *some* which ultimately gets stressed. First, interrogative *wh*-words also receive stress; this fact, taken in isolation, cannot be ascribed too much significance, for relative *wh*-words never receive stress, but their sources have in general been regarded as full-fledged NP's, which can be stressed. However, the stress correlation, together with the second fact I shall mention, does carry weight. Thus, stressed *some* has been described as 'specific,' whole unstressed *some* has been described as 'non-specific'. Specificity is a complicated notion, much more so than it had initially been thought, and must be defined with respect to (often proper) subparts of sentences (Dean-Fodor, oral communication). Thus, in (2.49a), *someone* is specific with respect to *John*, but not with respect to the speaker of the sentence.

(2.49) a. John thought that someone was in the building.
b. John wondered who was in the building.

Similarly, in (2.49b), *who* is specific with respect to John rather than with respect to the speaker of the sentence. The hypothesis that *wh*-words are specific can explain the fact that (2.49c) is non tautologous, for (2.49d) is not tautologous either, but (2.49e) is.

(2.49) c. Where is John?
d. John is sómewhere.
e. John is sŏmewhere.

Making the notion 'specificity' precise is a task far beyond the scope of this discussion, but if it can be argued that interrogative *wh*-words are always specific with respect to at least some clause, this will be sufficient motivation for regarding such words as derived from specific *some*, which will account in a straightforward fashion for (2.48).

With respect to (2.41c), I can report that I have presented various native informants with a more complete version of (2.43), which I render below as (2.50), and that none accepted (or had heard speakers who accepted) any of the sentences in (2.50) with the *j* or *k* linkages.

(2.50) a. Whose$_i$ father$_j$'s brother$_k$ did $\left\{\begin{matrix}\text{his}_i\\ \text{his}_j\\ \text{his}_k\end{matrix}\right\}$ sister criticize?

b. $\left\{\begin{matrix}\text{His}_i\\ \text{His}_j\\ \text{His}_k\end{matrix}\right\}$ sister criticized someone$_i$'s father$_j$'s brother$_k$.

c. The brother$_k$ of the father$_j$ of whom$_i$ did $\left\{\begin{matrix}\text{his}_i\\ \text{his}_j\\ \text{his}_k\end{matrix}\right\}$ sister criticize?

(2.50) d. $\left\{\begin{array}{l}\text{His}_i\\ \text{His}_j\\ \text{His}_k\end{array}\right\}$ sister criticized the brother$_k$ of the father$_j$ of someone$_i$.

Until more data can be gathered which corroborate Postal's, I think we can regard his data as suspicious and far-fetched. But even if we accept them, Postal's case is not yet established. Indeed, Postal discusses only (2.50a) and (2.50d), which he assumes are transformationally related, a not at all obvious point (and not a universally accepted one). But even conceding this point, we cannot be sure that crossing over a coreferent is responsible for the unacceptability of (2.50a) with the *j* and *k* linkages, so long as Postal does not tell us what his judgments are about (2.50b). If he rejects it (which is suggested by his failure to discuss it), the unacceptability of (2.50a) might be due not to cross-over, but to the preposing of genitives.

With respect to (2.41d), I have also been unable to find anyone willing to agree with (2.44). In fact, every informant felt that the b-sentence is better than the a-one, although both were judged acceptable. Postal himself points out that the unacceptability of (2.44b) is dialect-restricted, and I would guess that the dialect in question constitutes a definite minority. In any event, the existence of speakers who accept (2.43) or (2.44) does not constitute a difficulty for my proposal, for it is perfectly possible that certain language learners may reinterpret certain grammatical facts, when the crucial counterevidence is fairly rare and may not be encountered at the pertinent moment. Thus, rather than extracting the generalization that questions are ill-formed when their sources are ill-formed, a learner may infer that questions are ill-formed when there are coreferents of the question-word between its surface and deep positions. Such reinterpretation of data is perfectly possible when the facts do not unambiguously force some unique interpretation; for example, Carden cites the case of a family in which every member spoke a different 'dialect' with respect to quantifiers and negation. Therefore, although certain speakers may have reinterpreted the cross-over restrictions on questions as purely formal ones, it is still the case that these restrictions can be given a *natural explanation* for a large number of speakers.

There remains to answer questions (2.35c) and (2.35d) and the answer to both is quite simple. Concerning (2.35c), the constraints are confined to mentioned NP's for Wh-Rel-Movement, because it is only in this case that there exists a choice between two synonymous transformational outputs. Thus, if a crossing NP displaced by Wh-Rel-Movement were not the one mentioned, i.e., were not a wh-word, the head of the relative clause would not be coreferential with it, and its crossed coreferent could not have been relativized. With respect to questions, where the constraints are due to general constraints on pronominalization involving stressed *some*, non-mentioned NP's are exempt, so long as they do not violate pronominalization constraints; for example, *him* is not prevented from referring to Charley either in (2.34b) or in its source, (2.34a). Finally, non-mentioned NP's are exempt in the case of Y-Movement, because the mentioned NP is different from the coreferent of the unmentioned one, and the sentence is different in focus from the one that would result if the crossed coreferent were topicalized. Thus, the pertinent output is not ruled out by (2.22), since the complication has not been 'unnecessary', and it is not ruled out by (2.23), since the purpose for which a more remote NP was chopped is a valid one.

As far as (2.22) is concerned, fronting is the only means of identifying the mentioned NP in Y-Movement, but not so for questions and relatives, where the mentioned NP is identified by the wh-feature. Therefore, even if Pied-Piping applies, it is still clear which element is questioned or relativized; but if Pied-Piping were applicable to Y-Movement, it would be impossible to know what is topicalized. If it can be argued that chopping a larger NP is perceptually more difficult than chopping a smaller one, the situation can be explained by (2.22), as the hearer will always assume that the smallest possible NP has been moved, unless he is given a clear counterindication in the form of the wh-feature.

As I have pointed out already, any overt feature identifying the mentioned NP should do, whether it is wh (as in questions and relatives) or heavy stress (as in Clefting) (see (2.33)).

2.1.3. This section is concerned with the complexity of interrupted behavior (brought up in the preceding section in connection with the cross-over restrictions on Relativization and Topicalization), and especially with two structural discontinuities that must be handled concomitantly, which, as argued in 2.0.1,,create a comparatively high degree of complexity. The center of the discussion is the pseudo-cleft construction in English, a topic of considerable interest in its own right. Numerous analyses have been proposed for pseudo-clefts; none of them is entirely satisfactory, but I shall argue that an analysis based on the 'extraction' of the pseudo-cleft phrase comes closer to being adequate than analyses of other types so far proposed. The analysis I defend is weakened to some extent by the use of some ad hoc formal apparatus, but I argue that this drawback does not invalidate the essential claim of the section, namely, that the pseudo-cleft phrase is a member of a structural discontinuity, and that pseudo-cleft sentences become perceptually complex when that discontinuity is combined with another one, such that the two are improperly nested. I also argue that the ad hoc apparatus I use is a consequence of the presently available descriptive models which do not allow a viable alternative to movement transformations in stating the existence of discontinuous components. Ultimately, the purpose of this section is to show that the low acceptability of concomitant double discontinuities is (at least in part) responsible for the phenomena which led to the postulation of Ross' Right Roof Constraint (see (2.48) in section 1.2.4). Movements to the right may create several types of perceptual difficulty, and Ross, on the basis of a number of cases exhibiting grammatization concluded that the RRC is a purely formal constraint on grammars. Basing himself on Ross' claim, Bach (1971) attempted to provide a principled explanation for the observation that SOV languages do not seem to have a question-movement rule. Bach made the general assumption that syntactic categories and major transformations are identifiable across languages, as well as the following four specific assumptions about question-movement rules:

(2.51) a. The question phrase is based on an indefinite pro-form;
b. The question movement rule is unbounded;
c. The movement is to the front of the sentence or toward a governing verb (in fact, in the performative analysis, which Bach accepts, the movement is always toward a governing verb);
d. The RRC.

Assumptions (2.51b)-(2.51d) above can account for the absence of question-movements in a large number of derivations in SOV languages. Indeed, if the movement must proceed in one swoop, which is consistent with (2.51b), rather than by cyclic 'hopping', and if the movement is toward a governing verb which is necessarily to the right of the phrase to be moved, it follows that there will be a large number of derivations in which the question phrase will have to move to the right and outside the dominance of the sentence node which most immediately dominates it. Since such movements are precluded by the RRC, (2.51) successfully accounts for the lack of question phrase movements in such cases.

However, as illustrated in (2.52) with a Japanese example, question-movements never occur in SOV languages, not even when the question phrase belongs to a clause immediately embedded to the governing verb and when question movements would not violate the RRC.

(2.52) a. Watakusi-wa dare-ga pan o tabeta no ka to kiita.
I who bread ate asked
'I asked who ate the bread.'
b. *Watakusi-wa pan o tabeta no ka dare-ga to kiita.

To account for the ungrammaticality of (2.52b), Bach makes the following additional assumption:

(2.53) Assumptions like (a) to (d) [in (2.51)] are assumptions constitutive of possible transformational components, rather than possible derivations.

Let us call the position expressed in (2.53) the Strong Transformational Position, and the position upheld in Chomsky (1962, 1971) and Ross (1967) the Weak-Transformational Position. The latter claims that a transformation is blocked only when its application violates some constraint, the former claims that if a transformation violates a constraint in some derivations, then it can never apply in any derivation. It is clear that (2.53) can account for the lack of a question-movement rule in SOV languages (or for the fact that such a rule, if it exists, is always blocked), but it is also clear that the Strong Transformational Position cannot be correct in general, for practically every transformation I can think of is subject to some constraint, and the position in question would in fact prevent all transformations from ever applying.

Bach is not explicit on which constraints he wants to be subject to (2.53), but we can be reasonably sure that he does not regard all transformations as strongly constrained in the sense of (2.53); he was undoubtedly aware that other constraints of Ross', such as the CNPC or CSC, sometimes block Relativization or Question-Movement, but do not, for that matter, push these rules out of the grammar. Since Bach does not elaborate on this matter, I will interpret his phrase 'assumptions like (a) to (d)' in (2.53) as 'the assumptions (a) to (d)'; that is, I shall assume that he only claimed the RRC to be a strong transformational constraint.

There is clearly something unsatisfactory about postulating strong and weak transformational constraints without some principle able to explain why certain constraints are strong rather than weak. I will attempt to show in this section that the RRC is neither a strong nor a weak transformational constraint, but essentially a behavioral one. The fact that there are apparent strong transformational constraints need mean no more than that grammatized perceptual limitations have been generalized.

For my claim concerning the RRC to be empirically verifiable, it is necessary to find cases in which the pertinent restrictions are not grammatized (and, preferably, some plausible justification for grammatization in the cases in which it has occurred). Specifically, we must find at least one instance in which elements are allowed to move indefinitely far to the right, but in which the result varies in acceptability according to the values of some (set of) parameter(s). I wish to claim that there are at least two cases in point in English: (a) the rule which moves pseudo-cleft phrases (following Emmonds (1969a), I shall call it Focus Placement) and whose discussion forms the bulk of this section, and (b) the rule of Coordination-Reduction from left to right, which is briefly touched upon in note 19 (for an extensive defense of this rule, see Hankamer, 1971). As I have pointed out already, the extraction analysis involves an undesirable formal feature, namely, the postulation of empty nodes in underlying representation; such empty nodes are no great embarrassment for the Extended Standard Theory (Emmonds, for example, uses empty nodes quite freely, and, I shall argue, too freely), but they have no place in Generative Semantics which claims that only well-formed or meaningful semantic material has a legitimate place in underlying representation. I will attempt to show, however (in this section and in Appendix One), that my analysis, while postulating semantically irrelevant material, does not postulate ill-formed material in underlying representation; in addition, I will suggest a way of severely limiting the use of empty nodes. On the other hand, the analysis proposed by such generative semanticists as Ross or Bach does involve an incoherent underlying representation, and must therefore be judged inferior to mine. In fact, I do not know of any analysis so far formulated which enable both a semantically justifiable underlying representation and an observationally adequate statement of the pertinent syntactic generalizations; therefore, until a more adequate analysis becomes possible, I propose the provisional adoption of the account which follows.

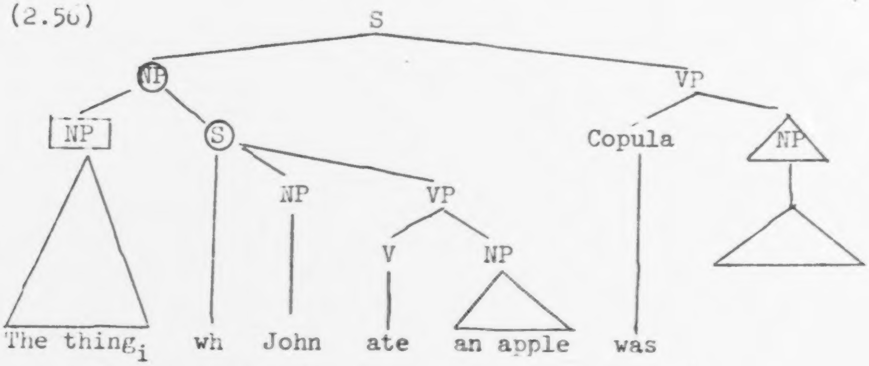
Consider the pseudo-cleft sentence in (2.54). Let us call (2.54b) the non-pseudo-cleft counterpart of (2.54a), and the underscored constituent, in either (2.54a) or (2.54b), the focus.

(2.54) a. What John ate was an apple.
b. John ate an ápple.

I make the following assumptions about the pseudo-cleft construction:

(2.55) a. The source of a pseudo-cleft sentence includes its non-pseudo-clefted counterpart with relative clause status. Specifically, I propose that a remote representation of (2.54a) is roughly (2.56).

(2.56)



Presupposition: John ate something$_i$.

The presupposition is the non-pseudo-cleft counterpart, except that the focus is replaced by a pro-form. The pro-form in the presupposition is the antecedent of the head of the relative clause in the underlying representation. This explains why the pro-form is indefinite and the head of the relative clause is definite.[10]

(2.55) b. (2.54a) is derived from (2.56) by
(i) placing wh on the focus;
(ii) copying the focus into the empty Predicate by a rule of Focus-Placement, such that a pro-form is left behind;
(iii) applying the rule of Relativization to attract the resulting wh pro-form to the head of the relative clause;
(iv) making optional morphophonemic adjustments.
c. The result of applying Focus-Placement is unacceptable
(i) if the instantiation of the essential variable[11] which intervenes between

---

[10] The problems involved in the semantic interpretation of pseudo-clefts are discussed in Appendix One.

the end of the sentence which most immediately contains the focus and the empty predicate is non-null, and

(ii) in proportion to the size and the nature of that variable.

---

[11] See Postal (1971: Ch. 13) for a discussion of this term.

First, I shall attempt to justify (2.55a) and (2.55b). All grammarians who have concerned themselves with pseudo-cleft constructions have recognized that the source of the pseudo-cleft must somehow incorporate its non-pseudo-cleft counterpart. This is so for at least two reasons:

(a) The selectional restrictions between the focus and its verb in the non-pseudo-cleft counterpart are the same as the selectional restrictions between the corresponding verb and the pseudo-cleft focus, as illustrated in (2.57);

(2.57) a. John broke a gláss.
b. What John broke was a glass.
c. *John broke a bóok.
d. *What John broke was a book.

(b) Certain elements appearing in position of neutralization, such as ever, any, and reflexive pronouns, are grammatical only if commanded by the appropriate neutralizing elements as shown in (2.58); however, this requirement need not be satisfied in pseudo-cleft constructions, provided that it is satisfied in the non-pseudo-cleft counterparts, as shown in (2.59); the neutralizing elements are underscored in both (2.58) and (2.59).

(2.58) a. I do not believe that Mary ever harmed anyone.
b. *I told the man who did not finish his dinner that Mary ever harmed anyone.
c. John admires himself.
d. *The girl who loves John also admires himself.

(2.59) a. What I cannot believe is that Mary ever harmed anyone.
b. I cannot believe that Mary ever harmed anyone.
c. What the missile did was destróy itself.
d. The missile destróyed itself.

The fact that the command-requirement with respect to neutralization is suspended in pseudo-cleft constructions, provided that it is satisfied in the non-pseudo-cleft counterpart, is quite general in English. (2.60)-(2.63) illustrate the point with Emphatic Reflexivization, Wh-Placement, and Sequence of Tenses.

(2.60) a. Jill believes that the letter was written by Ann and $\left\{\begin{array}{l}\text{herself}\\ \text{*himself}\end{array}\right\}$.

b. What Jill believes is that the letter was written by Ann and $\left\{\begin{array}{l}\text{herself}\\ \text{*himself}\end{array}\right\}$.

(2.61) a. John $\left\{\begin{array}{l}\text{told me}\\ \text{*believed}\end{array}\right\}$ where to go.

b. What John $\left\{\begin{array}{l}\text{told me}\\ \text{*believed}\end{array}\right\}$ was where to go.

(2.62) a. John $\left\{\begin{array}{l}\text{realized}\\ \text{*believed}\end{array}\right\}$ that he is stupid.

b. What John $\left\{\begin{array}{l}\text{realized}\\ \text{*believed}\end{array}\right\}$ is that he is stupid.

A representation like (2.56) is, however, not the only logically possible way of incorporating the non-pseudo-cleft counterpart into the underlying representation of pseudo-cleft constructions. (2.56), which is essentially the representation proposed in Chomsky (1970) and Akmajian (1970),[12] has been challenged by Peters and Bach (1968) on

---

[12] Actually, (2.56) differs in two respects from Chomsky's and Akmajian's proposals. First, it places the presupposition in underlying structure, while Chomsky and Akmajian regard it as a feature of surface structure. Second, and more important, Chomsky proposes that the head of the relative clause is a dummy it, while Akmajian proposes that it is sometimes a genuine NP and sometimes a dummy, as in (i) and (ii) respectively:

(i) The object that I saw was the house.
(ii) What I saw was the house.

My own view is that all free relatives such as (ii) have genuine heads at some remote level of representation. Akmajian's reason for postulating free relatives with dummy heads is that some free relatives do not have well-formed surface counterparts with genuine heads in his dialect; thus, (iii) is acceptable to him while (iv) is not.

(iii) Where John went was to Boston.
(iv) The place where John went was to Boston.

But surely this is no more than a surface phenomenon, because sentences like (iv) are acceptable to a great many speakers. In addition, headless relatives create much more serious problems, since the whole process of Relativization is based on there existing inside the relative clause a coreferent of the head; but how could a dummy be coreferential with anything? Notice also that overt forms of the dummy it, such as those associated with 'meterological' verbs two grounds:

(a) There is no natural way of carrying out step (2.55b.i) in such a way as to ensure that *wh* will be attached to only one NP.

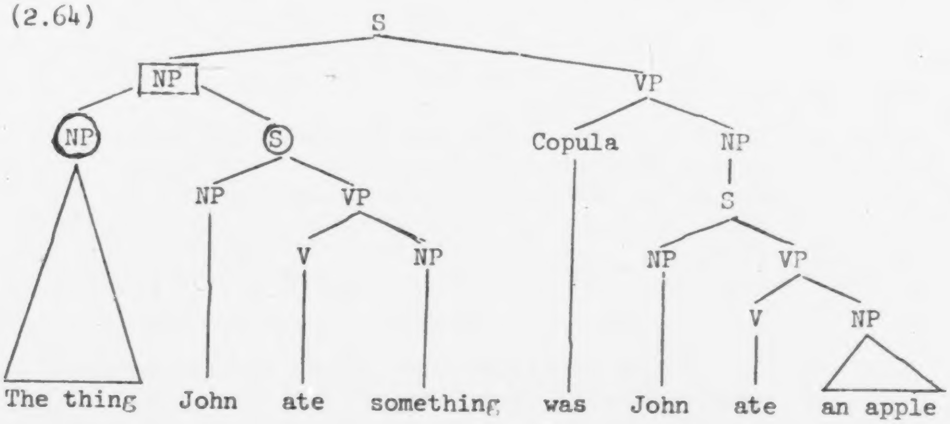
(b) Since the pseudo-clefted constituent cannot be selected in the base, it must be selected by an attachment transformation; however, the pertinent attachment transformation cannot be formulated, since the only NP's which can be pseudo-clefted are those which can be replaced by *something* with preservation of grammaticality, and the NP to be pseudo-clefted and the verb with which it participates in selectional restrictions may be arbitrarily far apart at the stage at which *wh*-attachment would have to apply.

Peters and Bach (1968) conclude that any analysis which involves the 'extraction' of the focus from a non-pseudo-cleft construction is inadequate, and propose an alternative analysis, originally suggested by Ross in personal communication:

(2.63) The source of a pseudo-cleft sentence contains the non-pseudo-cleft counterpart in predicate position in construction with the copula, while the subject consists of the NP *the thing* modified by the non-pseudo-cleft counterpart in which *something* has been substituted for the focus.

Specifically, the source of (2.54a) would be (2.64), rather than (2.56).

(2.64)



---

or resulting from Extraposition, cannot be modified by relative clauses, as shown in (v) and (vi):

(v) *It, which I dislike, is raining.
(vi) *It, which nobody yet knows, is certain that John has defected.

For a convincing argumentation that free relatives introduced by *where*, *when*, and *while* have meaningful heads in remote representation, see M. Geis (1970).

In this analysis, the rule of Focus-Placement is replaced by a rule which deletes the whole of the predicate except for the focus.

It is rather easy to subvert Peters and Bach's two objections to the extraction analysis and to show that their own proposal is inadequate on much more serious grounds.

Regarding their first objection, the difficulty they mention arises only because they propose to identify the focus by wh-placement, and because they propose to allow any NP to occur with wh-morphemes. This was understandable, in view of Chomsky's original proposal (see Footnote 12) which posited dummy-headed relative clauses for pseudo-cleft constructions. However, if we posit a meaningful head for the relative clause, the difficulty disappears, for the correct placement of wh reduces to a problem which must be solved for relative clauses in general.[13] One way in which this has been handled in the literature has been to place wh on a coreferent of the head inside the relative clause. (For ways of dealing with relative clauses which contain more than one coreferent of the head, see Postal (1971)). In the case of structures like (2.56), if wh-placement precedes Focus-Placement (which is the order given in (2.55b) because it is the one Peters and Bach criticize), wh can be placed on the constituent of the relative clause which has no identical counterpart in the presupposition. However, since the focus can be identified merely by comparison with the presupposition, Focus-Placement may precede Wh-Placement, and the latter can apply in perfectly straightforward fashion to the output of the former, i.e., to (A1.8) of Appendix One.

With regard to their second objection, it is simply incorrect that only NP's pronominalizable with something can be pseudo-clefted. Thus, a great many speakers accept not only (2.54a), but also the sentences in (2.65), which Peters and Bach reject.

(2.65) a. Where I met John was in Paris.
b. When I saw Harry was in the afternoon.
c. Why John left was to catch a plane.
d. ?Who Nixon chose was Agnew.

Clearly, however, the non-pseudo-cleft counterparts of (2.65a)-(2.65b) do not tolerate the replacement of the focus with something, therefore the Peters-Bach proposal fails to generate a large number of pseudo-cleft sentences in a large number of dialects. More important still, the Peters-Bach proposal fails to account for pseudo-cleft sentences like (2.66),[14] which are undoubtedly acceptable in all dialects,

---

[13] I am grateful to A. Zwicky for pointing this out to me.

[14] That sentences like (2.66) are instances of the pseudo-cleft construction follows from the fact that they exhibit the two properties typical of pseudo-cleft sentences which were mentioned at the beginning of this section. Indeed,

(a) The selectional facts exactly parallel those found in ordinary pseudo-cleft sentences, as can be seen by comparing (i) and (2.57) respectively.

(i) a. John broke a gláss.
b. The object which John broke was a glass.
c. *John broke a bóok.
d. *The object which John broke was a book.

(b) The predicate can contain *any-* or *ever*-words when the latter are *not* commanded by the neutralizing elements in surface structure, although this is impossible in general, as shown in (ii).

(ii) a. The insinuation which nobody could believe was that Mary had ever been involved with any shady characters.
b. *The insinuation which nobody could believe suggested that Mary had ever been involved with any shady characters.

In addition, sentences like (2.66) are subject to two constraints typical of pseudo-clefted sentences which are discussed below in the text, namely, the CNPC and the RRC", as can be seen by comparing (2.64) and (2.74) with (iii) and (iv) below respectively:

(iii) a. I suspect the idea which tortures Bill is that Mary loves John.
b. *It's Bill who I suspect the idea which tortures is that Mary loves John.
c. *It's John who I suspect the idea which tortures Bill is that Mary loves.
(iv) ?*The drink I told you John had at 3 o'clock at 4 o'clock is whiskey.

---

including theirs.

(2.66) a. The girl who just left is Mary.
b. The one who Mary admires most is herself.

Therefore, the Peters-Bach proposal is inadequate on observational grounds. The way to prevent the generation of the sentences in (2.65) in the Peters-Bach dialect is clearly not by limiting pseudo-clefted constituents to positions which may be occupied by *something*, but rather by blocking some of the morphophonemic rules required by (2.55b.iv) under the appropriate circumstances. Thus, while rules like *the place at which* ⇒ *where, the time at which* ⇒ *when, the manner in which* ⇒ *how*, etc. (which are independently needed for adverbial clauses in non-pseudo-cleft constructions, as shown in Geis (1970)), may be allowed to operate freely in most dialects, only the rule *the thing which* ⇒ *what* would be allowed such freedom in the Peters-Bach dialect, while the remainder would have to be blocked in pseudo-cleft constructions (presumably by restricting them to non-subject position).

Thus, neither of Peters and Bach's objections to an extracting analysis really go through. On the other hand, the alternative like (2.64) for (2.54a), must be rejected for the following two reasons (a third reason is provided at the end of this section):

(a) The most obvious defect of (2.64) is that it is semantically incoherent. Indeed, the meaning of (2.54a) is that the object which John is supposed to have eaten was an apple; it most certainly is not that the object in question was the preposition John ate an apple.[15]

---

[15] In personal communication, Ross has attempted to avoid the charge of meaninglessness by proposing that the subject of pseudo-cleft constructions is not a free relative, but an interrogative clause. Thus, the underlying representation of (2.54a) would no longer be (2.64), but something like (i).

(i) The answer to the question 'What did John eat?'
is 'John ate an apple.'

While (i) is meaningful, it cannot be correct for two reasons:

(a) it does not allow the generation of pseudo-cleft sentences with non-reduced heads in surface structure, such as (2.66).

(b) there is overwhelming syntactic evidence that the subject clauses of pseudo-cleft sentences are free relatives and not interrogatives. A large number of tests for discriminating between the two types of constructions was proposed in C. Baker (1968), and they all suggest that pseudo-cleft constructions involve free relatives. This is illustrated below with a few tests:

(1) Relative pronouns, unlike interrogative ones, cannot be stressed:

(ii) *Whát John ate was an apple.

(2) Subject interrogative clauses require singular verb agreement, while pseudo-cleft subject clauses do not:

(iii) a. Who walked into the room $\left\{\begin{matrix}\text{was}\\ \text{*were}\end{matrix}\right\}$ clear.

b. What copulated $\left\{\begin{matrix}\text{*was}\\ \text{were}\end{matrix}\right\}$ the dog and the bitch.

(3) Relative pronouns, unlike interrogative ones, cannot be modified by else:

(iv) a. What else you saw in the room is the question.
b. *What else you saw in the room is the apple.

(4) Interrogative clauses can contain more than one wh-word, while relative clauses cannot:

(v) a. What happened to whom isn't clear.
b. *What I saw when was Bill in the morning.

More recently, in Ross (1971), the Bach-Peters proposal is defended on the grounds that (vi) is a possible sentence in certain dialects.

(vi) What I ate is I ate an apple.

However, the fact that (vi) is a possible surface structure does not make it a coherent semantic representation. All the existence of (vi) means is that the whole circled S-node may be 'dragged along' by Focus-Placement in the pertinent dialects. If (2.64) is the underlying representation of (2.54a), its semantic interpretation can only be given by an interpretative rule which applies to the output of deletion in the standard dialect and to surface structure in the dialects which tolerate (vi). It is rather strange that Bach and Ross, who have consistently supported Generative Semantics in their work, seem to advocate an interpretivist approach in just this one instance.

(b) Neither elements of the subject nor elements of the predicate can be moved out of pseudo-cleft constructions, as shown in (2.67).

(2.67) a. I suspect what Mary told you Bill wants is to kill his mother.
b. *It's you who I suspect what Mary told Bill wants is to kill his mother.
c. *It's his mother who I suspect what Mary told you Bill wants is to kill.

Given an analysis of pseudo-cleft constructions as in (2.64), (2.67b) can be excluded in a principled way by invoking the Complex Noun Phrase Constraint (see discussion in 1.2.6). However, given a deletion analysis, or Postal's analysis described in Appendix One, there is no general principle which could exclude (2.67c), for elements can in general migrate quite freely from predicate position, and an ad hoc condition would have to be specified for blocking movements out of pseudo-clefted constituents. To see this more clearly, contrast (2.68a), which is ambiguous between a pseudo-cleft and a non-pseudo-cleft reading, and (2.68b), where the non-pseudo-cleft reading alone is possible:

(2.68) a. The discovery John announced was a proof of Mary's defection.
b. It's Mary's defection which the discovery John announced was a proof of.

On the other hand, given an analysis of pseudo-cleft constructions as in (2.56), both (2.67b) and (2.67c) can be excluded by the CNPC,[16]

---

[16] I have found that there are speakers who can question (but not relativize) the *whole focus*, even though *elements of the focus* can be neither questioned nor relativized. That is, the paradigm in (i) holds for those speakers.

(i) a. What do you think that what John ate was?
b. *The apple which you think that what John ate was had been bought in Canada.

If we accept Akmajian's (1970) claim that cleft sentences are derived from pseudo-cleft ones, the distinction between questioning and relativizing the focus can be seen more clearly in (ii).

(ii) a. What is it that John ate?
b. *The apple which it is that John ate had been bought in Canada.

The point I wish to make is that the acceptability of (iia) and/or (ia) is no counterexample to my claim that the focus originates inside a complex NP. The questioned constituent is the triangled NP in (2.56), which does not originate inside a complex NP, and is therefore free to move.

The reason (ib) and (iib) are bad is that foci are necessarily stressed, while relative pronouns (unlike interrogative ones) are necessarily unstressed.

In fact, I believe that a deeper explanation is possible here: a cleft or pseudo-cleft constituent is the comment of the sentence (i.e., the new information that it imparts), as opposed to the remainder of the cleft or pseudo-cleft construction, which constitutes the topic (or old information, since it is presupposed); as new information usually receives stress, the fact that the cleft or pseudo-cleft phrase constitutes new information is (at least in part) responsible for the fact that it is stressed. Questions are similar to cleft or pseudo-cleft sentences (this fact has been noticed before); thus, what John saw was a dog presupposes John saw something, while in the case of What did John see? the speaker believes John saw something to be true. Question wh-words are indicators of where the comment should be supplied in the 'expected answer' (i.e., in a declarative sentence identical with the question in all respects, except that the wh-word is replaced by a more informative phrase); therefore, it is not surprising that cleft or pseudo-cleft constituents should be able to be questioned, since the expected answer necessarily has the comment in the position of the wh-word.

On the other hand, a relative wh-pronoun is the topic of the relative clause, while the remainder of the clause is its comment (this is probably why relative pronouns are always unstressed). Therefore, constructions like (ib) or (iib) imply that the relative pronoun is both a topic and a comment, that is, both old and new information, an obvious impossibility. Notice that we are dealing here not with perceptual conflict, but with semantic ill-formedness, for it is possible to infer a contradiction from (ib) or (iib), namely, that which is simultaneously new and old information; consequently, I would not expect (ib) or (iib) to be possible sentences in any dialect of English (or in any language in which identical conditions are allowed by the grammar).

---

since the pseudo-clefted constituent, which contains mother in underlying representation, originates inside a complex NP (such as the circled one in (2.56)).

Notice that sentences like (2.67b) and (2.67c) provide independent support for the claim I made in footnote 12 that all free relatives should be represented with genuine (i.e., lexical) heads in remote structure, for unless we make this assumption, neither (2.67b) nor (2.67c) can be ruled out by the CNPC, and both would necessitate ad hoc conditions.

I take it that the arguments given so far demonstrate the superiority of (2.55a) and (2.55b) over the Ross-Peters-Bach analysis of pseudo-cleft constructions. Now I address myself to the main point at issue, namely, the interplay of the rule of Focus-Placement and the RRC.

Let us first notice that Focus-Placement must be an unbounded rule, as suggested by (2.69).

(2.69) What I forgot to tell you that Mary insinuated that Bill was prepared to instruct Greg to tell the FBI was that you intended to defect to Russia.

Therefore, given Ross' formulation of the RRC, Focus-Placement should be an impossible rule, and this argument has in fact been used by Ross against all analyses of pseudo-cleft constructions involving extraction. Chomsky (1971) has replied to this objection by offering a slightly modified version of the RRC, which we shall call the RRC', and which stipulates that unbounded rightwards movements are ungrammatical only if the essential variable they cross is non-null (this is equivalent to my (2.55c.i)). The RRC' undoubtedly rises higher in observational adequacy than the RRC, for it makes it possible both to derive sentences like (2.69) through extraction, an analysis which, as we saw, is to be preferred to the Ross-Bach-Peters one on independent grounds, and to block all the sentences which the RRC can block. To see this, we need only examine some of the cases which supplied the motivation for the RRC and recognize that the ungrammaticality of the crucial constructions may be ambiguously attributed to the RRC or the the RRC'. Thus, consider the rules of Particle-Movement, Extraposition, Extraposition-from-NP and Right-Dislocation exemplified in (2.70)-(2.73) respectively; the a-sentences arise when the corresponding rule has not applied, the b-sentences when it has applied without crossing the boundary of the sentence immediately containing the moved constituent, and the c-sentences when it has applied and crossed the boundary in question. Notice that in each of the c-sentences, the moved constituent has crossed not merely an essential variable, but a non-null one.

(2.70) a. You testified that my brother had looked over the desk, although you knew that your testimony could send him to jail.

b. You testified that my brother had looked the desk over, although you knew that your testimony could send him to jail.

(2.70) c. *You testified that my brother had looked the desk, although you knew that your testimony could send him to jail, over.

(2.71) a. That that John has disappeared is odd is surprising.
b. That it is odd that John has disappeared is surprising.
c. *That it is odd is surprising that John has disappeared.

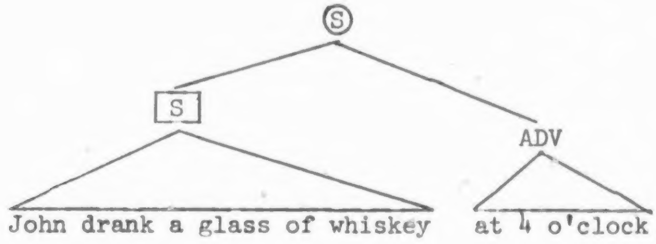
(2.72) a. That the claim that John is a traitor is unbelievable is odd.
b. That the claim is unbelievable that John is a traitor is odd.
c. *That the claim is unbelievable is odd that John is a traitor.

(2.73) a. That you have been annoying my sister is unbelievable.
b. That you have been annoying her, my sister, is unbelievable.
c. *That you have been annoying her is unbelievable, my sister.

Despite its greater appeal, the RRC' itself cannot be correct, for it wrongly predicts that (2.74) is ungrammatical. This is so, if we accept the claim, which I defended in (.14) in connection with objection (g) to the IC, that the structure of the non-pseudo-clefted counterpart of (2.74) is esentially (2.75).

(2.74) What John drank at 4 o'clock was a glass of whiskey.

(2.75)



Indeed, the focus *a glass of whiskey* must cross the non-null *at 4 o'clock* which lies outside the boundaries of the boxed S. On the other hand, we would be ill-advised to reject the RRC' completely, since it correctly predicts the unacceptability of (2.76)-(2.78).

(2.76) ??What I told you John believes in on Monday morning at 4 o'clock is God.

(2.77) ?*What I told you John drank at 3 o'clock on Monday morning at 4 o'clock is a glass of whiskey.

(2.78) *What I told you John drank in order to please his girl friend from Chicago at 3 o'clock on Monday morning at 4 o'clock is a glass of whiskey.

How can we preserve all the advantages of the RRC' and at the seme time account for the acceptability of (2.74) as well as for the gradually increasing unacceptability of (2.76)-(2.78)? I propose to do this by taking into account not only the fact of a non-null instantiation of a variable, but also the size and the nature of such an instantiation, as required by (2.55c.ii). (We shall call this proposal the RRC".)

With respect to size, it should be pointed out that sentences like (2.74) become increasingly awkward in proportion to the length and complexity of adverbials, as shown by (2.79)-(2.80).

(2.79) ?What John drank before our big fat neighbor living on Front Street and wearing a red beard came in is a glass of whiskey.

(2.80) ??What John drank before our big fat neighbor who lives in Chicago and whose only daughter had an unfortunate love affair came in is a glass of whiskey.

The awkwardness of sentences like (2.79) was ascribed by Bever to the already mentioned interruption principle (2.37):

(2.37) Discontinuous components are complex in proportion to the structural complexity of the intervening material.

(2.37) is a plausible principle, since, given an immediate memory with limited capacity, we may expect that retention of the first member of a discontinuity to be heard in proportion to the amount of processing that must go on before the remainder of that discontinuity is encountered and its first member can be discarded from immediate memory.

With respect to the nature of the instantiation of the pertinent variable, let us first notice that it must be taken into account in addition to considerations of size, since the size of the adverbial is the same in (2.74) and (2.76), but the two sentences differ sharply in acceptability. The crucial difference can be appreciated by comparing (2.81) (a slightly modified version of (2.74) to make it similar to (2.76) in all respects except the crucial one) and (2.76).

(2.81) What I believe John drank on Monday morning at 4 o'clock is whiskey.

In (2.76), but not in (2.81), the adverbial forms a constituent with material which precedes the sentence from which the pseudo-clefted constituent was extracted. I suggest that the unacceptability of (2.76) is due to the fact that two interrupted sentences are continued separately, so that a hearer has to figure out which continuation is to be associated with which previously interrupted sequence. I therefore propose the following tentative principle:

(2.82) Two interrupted sentences are perceptually complex
if one of them is at least in part internal
(i.e., center-embedded) to the other.

It is easy to see that (2.82) can account for the unacceptability of all the c-sentences in (2.70)-(2.73). For example, in (2.70c), the interrupted *that*-clause is internal to the matrix sentence which the *that*-clause itself interrupts; the difficulty results from the fact that both sentences have been interrupted after *desk*, and from the fact that the *although*-clause must be associated with *you testified*, while *over* must be associated with the *that*-clause.[17] An independent virtue of (2.82) is that it can account for the complexity of center-embedded sentences to a degree higher than one, such as (2.83).

(2.83) *The girl who the teacher that the school fired flunked cried.*

The triply- and doubly-underscored sentences are both interrupted, and the doubly underscored one is 'at least in part internal' to the triply-underscored one; as in the previously considered examples, the difficulty is mainly due to the fact that *flunked* and *cried* must be paired with the appropriate distinct sequences.[18], [19]

---

[17] The rules illustrated in (2.70)-(2.73) also provide evidence against both the RRC and the RRC' and in favor of the RRC", for the limitations on their scope are not in fact identical, as Ross thought. Thus, in structures like (2.75), the scope of Particle-Movement is indeed the boxed S, but the scope of Extraposition is the circled S, as shown in (i) and (ii) respectively.

(i) a. That John may die was announced at noon.
b. It was announced at noon that John may die.

(ii) a. John looked over the desk at noon.
b. John looked the desk over at noon.
c. *John looked the desk at noon over.

The restrictions on Particle-Movement are reminiscent of the situation found in connection with question phrase movements in Japanese: the RRC" has been both grammatized and strengthened. Extraposition provides an illustration of the intermediate situation in which the RRC" has been grammatized, without strengthening.

[18] My proposal is similar, but not identical, to the following explanation of the complexity of sentences like (2.83) proposed in Chomsky and Miller (1963) and Chomsky (1965):

(i) A device cannot interrupt a given procedure more
than once in order to use that very procedure.

I believe (2.82) to be superior to (i) in two respects:

(a) (2.82), but not (i), can account for both (2.83) and for (2.76)-(2.77) (as well as for (2.70)-(2.73c)). This is so because (i) depends crucially on their existing a *third sequence*, such as the singly-underscored one in (2.83), which must moreover be of the same kind as the two interrupted sequences, while (2.82) merely requires that two distinct interrupted sequences be *separately* continued;

(b) As Bever has pointed out, it seems ad hoc to propose that one interruption of a procedure by itself is perfectly acceptable, while two such interruptions are totally unacceptable; this criticism does not apply to (2.82), since it is irrelevant whether the two sequences are interrupted by a third one, as in (2.83), or by each other, as in (2.76). It is interesting to note that the alternative explanation proposed by Bever, which I reproduce as (ii) below, cannot explain all the facts about centerembedding either, even though it may be able to explain some of the facts.

(ii) A stimulus may not be perceived as simultaneously having two positions on the same classificatory dimension.

Indeed, Bever blames the unacceptability of (2.83) entirely on the 'double-function' of *flunked*; notice, however, that (iii) and (iv), in which the underscored elements also have double-function, are infinitely better than (2.77).

(iii) The buckles of *the collars* of these coats are strong.
(iv) John's *father's* brother is sick.

---

[19] Notice that the requirement that the two sentences be *separately* continued (which follows from (2.82), can explain the acceptability of sentences like (ii) below, which is derived from (i) through Coordination-Reduction, in violation of the RRC and of the RRC'.

(i) I told Mary that John bought the cake, and I also told her that he later ate the cake.
(ii) I told Mary that John bought, and I also told her that he later ate, the cake.

Given an analysis of the Conjunction-Reduction which copies *the cake* simultaneously out the two conjuncts and Chomsky adjoins it to the coordinate node, the copying of *the cake* out of the second conjunct violates the RRC and its copying out of the first conjunct violates the RRC'. However, the derivation of (ii) from (i) does not violate (2.82), since the second conjunct is not internal to the first; in fact, both discontinuities are resolved *simultaneously* by *the cake*, rather than *separately*, as in (2.76) or (2.83).

If the unacceptability of (2.76) can be attributed to (2.82), the greater unacceptability of (2.77) and (2.78) can be attributed to the combined effect of (2.82) and (2.37). Indeed, in (2.77) and (2.78) the instantiation of the pertinent variable is both longer and more complex than in (2.76), since it must be segmented into at least two parts, each of which must be paired with different sequences.[20]

---

[20] There are speakers who reject all the sentences in (2.76)-(2.78) indiscriminately. I consider that (2.82) has been grammatized in their dialects.

I take it that the arguments I have advanced in support of (2.37) and (2.82) have shown that the RRC" (i.e., (2.55c.i) and (2.55c.ii) with size and nature interpreted in the sense of (2.37) and (2.82) respectively) is superior to both the RRC and the RRC' as an account of the Right Roof Constraint. But now we have an additional reason for adopting (2.55a), (2.55b) and rejecting the Ross-Peters-Bach analysis. This is so, not because the former involves movement of the pseudo-clefted phrase (indeed, the behavioral position does not depend on there being a transformation involved--since (2.82), for example, is applicable to structures like (2.83), where the pertinent discontinuities do not result from the application of a transformation); rather this is so because, given a representation like (2.64), the first occurrence of John ate and an apple do not form a discontinuous component.[21] That

---

[21] My claim that the marginal status of sentences like (2.76) is due to the existence of a perceptually complex discontinuity makes it irrelevant--at the explanatory level--whether Focus-Placement proceeds in one swoop or by cyclic hopping.

---

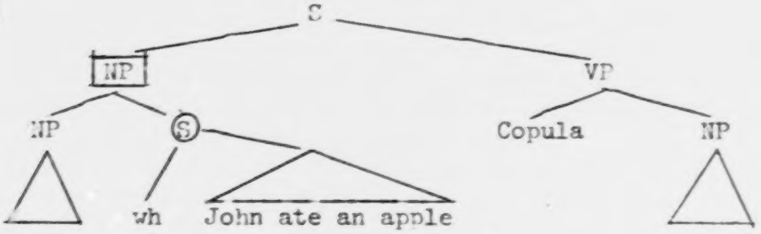
they should be represented as forming a discontinuous component is also shown by the difference in acceptability between (2.84a) and (2.84b), which differ only in that the former is a pseudo-cleft sentence while the latter is not.

(2.84) a.??What I said that John likes most in front of all the people who had assembled in front of his house is a bottle of good French wine.
b. What I said that John likes most in front of all the people who had assembled in front of his house is no concern of yours.

Before concluding this section, an evaluation of the merits and demerits of the particular analysis of pseudo-clefts I have proposed is in order.

The underlying representation I propose for (2.54a), namely, something like (2.85)--(See Appendix One), has the disadvantage of allowing empty nodes in underlying representation.

(2.85)



Presupposition: John ate something.

Empty nodes have been extensively used by Emmonds in his dissertation and in other papers in an effort to show that certain transformations are *structure preserving* (i.e., they do not create more structure). To give his hypothesis empirical content, Emmonds requires that empty nodes be allowed only where a non-empty constituent of the same type can be generated by the base rules (empty nodes carry categorical labels and are subject to subcategorization). The empirical test of the hypothesis is that constituents should not be movable to a position where a constituent of the same type cannot be generated in the base. Basing himself on this assumption, Emmonds (1969b) presents a rather ingenious argument for *there* insertion into an empty node. He argues that *be* usually require a subject and a complement, as in *my uncle is very old*, but that with indefinite subjects, it looks like the complement may be missing, as in **a huge fire was*, in which case *there-insertion* is obligatory and yields *there was a huge fire*. Emmonds proposes a 'simpler' solution for the latter case: *a huge fire* is the complement rather than the subject, the subject node is generated empty by the base, and *there* is inserted into it. The problem with this analysis is that it implies that the form *be* occurring in these two examples is *the same* verb. However, it is clear that in the former case, *be* is a two-place predicate asserting a relation between my uncle and the property of being very old, while in the latter, it is a one-place predicate asserting the existence of a huge fire. Consequently, the generation of an empty node for *there* is quite ad hoc in this case and with no semantic justification (I have criticized Emmonds' principle for its unnaturalness, while assuming it can do the job it is supposed to do; however, Emmonds' treatment of *there*-Insertion is inadequate on observational grounds, for while the one-place predicate *be* usually requires *there*-Insertion, there are a few expressions in which it does not, such as *God is*, or *I think, therefore I am*, which Emmonds would be forced to claim have an empty node in surface structure and are therefore ill-formed).

A more restricted proposal involving empty nodes would require that empty nodes should be allowed only if they, together with the non-empty arguments, add up exactly to the number of arguments that the pertinent predicate allows on logical grounds. This requirement is not violated in (2.85), for *be* is a two-place predicate in pseudo-cleft construction, and empty nodes may be allowed with the provision stated above, when their presence has considerable independent advantages. I review and evaluate these advantages below.

First, the fact that the pseudo-cleft phrase originates inside the circled S-node in (2.85) enables us to account for the fact that its selectional restrictions are determined by the verb of which it is an argument in the non-pseudo-cleft counterpart. However, J. Gundel has pointed out to me that an alternative account is possible. Thus, the copy which the extracted phrase was claimed to leave behind is coreferential with the head of the relative clause; as the head can enter into the same selectional restrictions as the relative pronoun it attracts (since they are coreferential) and as the pseudo-cleft construction asserts coreferentiality between the referent of the head (further identified by the relative clause) and the more specific predicate NP, it is not surprising that the latter can take part in the same selectional restrictions as the less specific head of the relative clause, and therefore of the copy which is ultimately relativized. Therefore, extraction is not the only possible solution here.

Second, the extraction analysis accounts for the island character of cleft and pseudo-cleft phrases. I do not see what alternatives are available, but P. Schachter (personal communication) has pointed out to me that some non-pseudo-cleft copular constructions exhibit the same properties. Thus, the doubly-underscored element in *the discovery was that John likes Mary* cannot be questioned, as **who do you think the discovery was that John likes* is just as bad as (2.67c). True, most copular subjects which disallow movements out of the predicate are themselves paraphrasable by a sentence (thus, a more remote representation which Generative Semantics would assign to *the discovery was that John likes Mary* might be something like *the thing which was discovered was that John likes Mary*). However, the claim that **who did you say the discovery was that John likes* is bad by the CNPC would be somewhat contrived.

Third, the extraction analysis makes it possible to say that the pseudo-cleft phrase and some sentence in the subject clause form a discontinuous component. It may legitimately be asked whether extraction is the only way of expressing a discontinuity. Personally, I do not know of a viable alternative, but I would undoubtedly welcome one which would do the same work as my analysis and would not need to employ empty nodes. I would, however, like to point out that movement or copying transformations have *already* been used in the literature to express the fact that a structural discontinuity exists, even though those cases involved exactly the same kind of difficulty as my account of pseudo-clefts with respect to underlying representation. For example, Ross (1967) posited a rule of Right-Dislocation to account for sentences containing after-thoughts, such as (2.73b); specifically, Ross proposed that (2.73b) is derived from its 'non-dislocated counterpart', (2.73a). That right-dislocated structures involve a discontinuity is strongly suggested by the fact that they obey the RRC (see (2.73c)), as well as by the fact--not mentioned by Ross--that they allow (apparent) violations of the command-requirement, as far as feature-changing is concerned, in exactly the same way as pseudo-clefts do. Thus, in parallel manner to sentences like (2.59a) or (2.59c) there are sentences like *I don't believe it, that Mary ever harmed anyone* or *she did it, compromise herself irrevocably* respectively.

The difficulty is that the main clause and the afterthought can be separated by such items as *I mean*, *namely*, *that is*, etc., and it is not possible to regard *I don't believe it, namely, that Mary ever harmed anyone* to be derived from **I don't believe that Mary ever harmed anyone, namely*, for such a representation is semantically ill-formed, as it makes the implicit claim that *namely* is a one-place predicate. It seems to me that it would be more acceptable to regard the source of sentences containing *namely* as having the latter as a two-place predicate, one argument being the 'non-dislocated counterpart' and the other being empty; the dislocated phrase can then be copied into the empty node by Right-Dislocation, just as the pseudo-cleft phrase is copied into an empty node by Focus-Placement. As in the case of pseudo-clefts, empty nodes strike me as more tolerable than no nodes at all.

Fourth, the extraction analysis makes possible a straightforward account of the behavior of all feature-changing processes, as shown in (2.58)-(2.62), and this seems to be the strongest argument in its favor. The fact that all the feature-changing processes behave as if the neutralizee were commanded by the neutralizer cannot be accounted for by an analysis which takes the underlying form of pseudo-clefts to be essentially their surface form; it can, of course, be accounted for by the Peters-Bach-Ross analysis, but that analysis postulates an incoherent underlying structure, while mine merely involves unnecessary material. Consequently, until a superior analysis of pseudo-clefts is proposed, I believe the extraction one can stand.

Regardless of these theoretical considerations, my fundamental concern in this section has been to show that the interplay of Focus-Placement and the RRC" can be handled adequately only within the Behavioral Position, which predicts that the derivations in which the RRC" is violated can be rated on a gradual scale of unacceptability in a principled way. The Strong Transformational Position fails with respect to Focus-Placement, for it predicts that there are no pseudo-cleft sentences in English. The Weak Transformational Position, which predicts that the derivations in which a constraint is violated are all ungrammatical, seems to be supported by those dialects which indiscriminately reject (2.76)-(2.78) (see footnote 20), but cannot account for the majority dialects. On the other hand, the Behavioral Position can account for the dialects described in footnote 21, by assuming the grammatization of (2.82).[22]

---

[22] As I pointed out in the Introduction, the problem of grammatization cannot be seriously considered before the significance of all the pertinent variables is known. The problem might be even more complicated, for grammatization may depend not only on the cognitive endowment of language learners, but also on their motivations; in other words, we may need not only an explicit theory of cognition, but also explicit theories of thought, behavior, and human interaction--still a remote goal at the moment. If, however, I may allow myself to speculate on some motivational factors with respect to the cases

considered in this section, it seems to me that a not implausible guess as to the failure of the RRC" to grammatize in pseudo-cleft constructions is that the likelihood of occurrence of pseudo-cleft sentences which violate the RRC" as compared with those which do not is relatively small. Moreover, sentences like (2.76)-(2.78) have stylistic variants which do not violate the RRC"; for example, (2.78) can be rendered by the quite acceptable (i).

(i) What I told you on Monday morning at 4 o'clock that John drank at 3 o'clock in order to please his girlfriend from Chicago is a glass of whiskey.

In contrast, if question-movements had to occur in SOV languages, the RRC" would probably be violated in a large number of derivations, and consequently a large number of questions would be marginal or totally ungrammatical, since the material over which the question phrase would move would have to contain verbs (as shown in (ii)), as there are no stylistic variants in which these verbs can be fronted, for the verb is necessarily final in Japanese.

```
(ii) Watakusi-wa Biru-wa dare-ga suki ta to Mary-wa
        I        Bill     who    like       Mary
       shinjite iru to John-wa utagatte iru no ka to kitta.
       believes        John    suspects              asked
     'I asked who John suspects Mary believes Bill likes.'
```

One objection which could be raised against this account is that, so far, no SOV language has been found in which question-movements operate, and if the RRC" is indeed a perceptual principle, we may expect that there should be SOV languages in which question-movements operate freely from the sentence immediately below the question verb, and with increasing awkwardness as we question from deeper levels of embedding. While such an objection cannot be passed over lightly, it should be pointed out that the predicted likelihood of finding question-movements in an SOV language is very small, as there is really no pressure in this direction. Indeed, so long as the question phrase is overtly identifiable, there is no stringent need to move it. Thus, there are dialects of English in which question-movements are optional, that is, in which (iia) is interpretable as a genuine request for information; moreover, question verbs can govern more than one wh-word, only one of which is allowed to move and still sentences with more wh-words than question verbs are perfectly interpretable, as shown by (iiib).

(iii) a. You were talking to whom?
      b. I wonder who saw what.

The assumption that motivational criteria like the one proposed above can determine grammatization is supported by the behavior or Relativization in Japanese. This rule is subject to the CNPC, the CSC and the Crossover Constraint (Ross (1967)), unlike the Japanese rule of Question-Formation (analogous to wh-placement), which is subject to none of these constraints. As Relativization obeys most of the constraints on movement rules, we would expect it to obey the RRC", but it does not, as the acceptability of (iv) shows.

(iv) Biru-ga suki ta to Mary-ga shinjite iru to John-ga
Bill likes Mary believes John
utagatte iru onna no ko wa kiree da.
suspects girl pretty.
'The girl who John suspects Mary believes Bill likes is pretty.'

This suggests that constraints can be violated when the functional motivations for doing so are strong enough, for if Relativization were also subject to the RRC", there would hardly be any relative clauses in Japanese, and apparently a language cannot easily get by without relative clauses. (iv) also supports my hypothesis that the RRC" is a behavioral principle, for it is conceivable that a behavioral principle can be violated, but it is hardly possible to claim that a universal syntactic constraint sometimes holds and sometimes does not.

Concerning the rules illustrated in (2.70)-(2.73), we may consider some possible reasons for their grammatization. Notice that in the case of the pseudo-cleft construction and of Coordination-Reduction to the right, there are reasons for the dislodged constituent to rise to a certain height: the pseudo-cleft phrase must be a clause-mate of the head of the subject relative clause, and the reduced constituent must be lifted high enough to escape from the domination of the coordinating conjunction (and or or). No such motivations exist with respect to the rules exhibited in (2.70)-(2.73), and we find no movement even when the extracted constituents would cross a null-variable.

In addition to considerations of double-interruption, there may be additional factors contributing to the unacceptability of the c-sentences in (2.70)-(2.73). Thus, consider the case when the remainder of the sentence form which some element has been displaced to the right by a non-constituent, as in (2.70c). This would force the hearer to keep in mind a non-constituent, presumably a more difficult task than the retention of a constituent (this has, I believe, been shown experimentally); in contrast, since rules do not move non-constituents, unbounded movements to the left could never force the retention of a non-constituent, and this difference in the effects of movements might contribute to the limitation of rightward movements more severely than of left-ward movements.

On the other hand, Particle-Movement does not necessarily leave behind strings which look like non-constituents, as for example in *that you claim that Bill had seen Mary is odd through; in this situation we may reasonably assume that the closure strategy (2.8) increases the complexity of this sentence, for the occurrence of the new predicate is after Mary presumably constitutes a sufficient encouragement for the hearer to regard the strong Bill had seen Mary as a complete sentence. The same situation favoring erroneous closure exists in (2.72c), and probably even in (2.71c). With respect to (2.73c), its unacceptability may look puzzling, for antecedents (more correctly, postcedents) can usually come arbitrarily far after pronouns. However, we are not dealing here with an ordinary postcedent, but rather with a tag expressing an after-thought, like I mean, my sister. Notice that such tags are awkward when delayed even if they have not been lifted into a higher clause, as shown by the unacceptability of *you've been pursuing her, although I explicitly forbade you to get out of your room, (I mean,) my sister, in which the although clause is an inserted parenthetical. Contrast this with clauses moved by Extraposition, which allow the insertion of long patentheticals: It is surprising to me, although I might have expected it all along, that John has left with another woman. The general conclusion we may draw is that erroneous closure is much more likely when elements move to the right than to the left, for leftward displaced elements usually carry a distinguishing mark which alerts the hearer to the fact that a discontinuity exists. Such closure is, however, less likely in the highly characteristic pseudo-cleft construction than in sentences like those in (2.71)-(2.73).

2.2.0. This section discusses various 'minimal distance' principles which have been often claimed 'to exist in order to reduce the amount of ambiguity' that would otherwise occur in natural languages. I believe that, put in this way, the claim is misleading, since it suggests a teleological orientation of speakers aimed at minimizing ambiguity; if speakers really followed a strategy like 'reduce ambiguity whenever possible,' it is difficult to see why they should allow ambiguity at all. I prefer to think that there are principles which have the indirect effect of reducing ambiguity in certain cases. One such principle could be (2.22), which prompts speakers to select, in the absence of indications to the contrary, the simplest reading of an ambiguous utterance as the most likely; whenever grammatization occurs (marking the less likely reading as ill-formed), the result is a reduction in potential ambiguity. One (usually) grammatized instance was seen in connection with the crossover restrictions on Relativization and Topicalization; some non-grammatized cases are considered below.

It will also be argued that in cases of equal likelihood, ambiguity is possible only when the potential readings are, in a sense ill-understood at present, minimally 'distinct.' For example, the sentence *John has a dog, Mary has a cat, and Bill has one too is impossible, for the hearer has no basis on which to select one of the (at least) two potential antecedents of one; as no reading is more likely than the other, both are impossible.

Granted that reduction of ambiguity is not a primary goal of speakers but an indirect effect of independent interacting principles, it still looks 'intuitively reasonable' to expect that minimal distance might increase the likelihood of one out of two or more potential readings. In fact, a linear minimal distance principle might even appear to derive experimental support from such well-attested facts of perception as the recency and primacy effects (which say that the first and last elements of a sequence are more easily remembered than the remaining ones). Such a claim would, however, be fallacious, for the closest *pertinent* element (for example, the most recent of a number of possible antecedents) need not be the *last element heard before the pro-form*; therefore, the recency effect does not support a linear minimal distance principle *in general*.[23] In fact, I will attempt to show below that whenever a minimal distance principle seems to be operative, there exists a more plausible alternative explanation.

2.2.1.1. One of the most simplistic proposals involving a minimal distance principle can be found in some traditional textbooks which lay down that if a pronoun has several possible antecedents, the actual antecedent will be the closest to the pronoun, in terms of linear distance. That his claim is empirically false is shown by the ambiguity of (2.86).

(2.87) John told Bill that he had become a father.

(2.86) shows clearly that the prescriptive pronouncement referred to above is not a rule of English. In fact, the choice of the closer antecedent is not even more likely, which suggests that the choice of a remote antecedent creates no perceptual problems in principle; in other words, the antecedent pronoun relation does not seem to be subject to the usual perceptual constraints on discontinuities (such as (2.37)).

That choosing the more remote antecedent in the case like (2.86) is not perceptually complex is also suggested by the fact that there is apparently no limit to the amount of discourse that can separate a pronoun from its antecedent, so long as the topic of the discussion has not changed radically in-between. Imagine, for example, one girl asking another how she had spent her honeymoon with John in Paris, and the newly married one launching into a description of the beauties of Paris, of the magnificence of the Bois-de-Boulogne, of the splendor of the Jardin-des-Tuileries, which she would conclude with the sentence: *and in front of the statue of Louis XIV at Versallies, he kissed me passionately*. It seems to me that even if the account contained no reference to John, there is still no difficulty whatsoever in associating *he* in the above underlined sentence with *John*. The hypothesis that establishing a connection between two NP's is *not* complex in proportion to the amount of intervening discourse will play an important part in my attempt to account for what I called in Chapter One the Dichotomous Behavior Principle.

---

[23] There is one case in which the recency or primacy effect seems to be involved. Thus, a question like *would you like milk, tea, or coffee*? can be answered *yes* in some languages, with the meaning of either *milk* or *coffee* (but not both for the same language). The meaning *milk* seems to be possible in a number of American dialects, while the meaning *coffee* is possible in Burmese (I owe this piece of information to A. Zwicky); I do not know of any language in *which* the answer *yes* to the above question can mean *tea*, and if none in effect exist, this would follow from the non-existence of a strategy which selects a medial element in a sequence.

2.2.1.2. My claim that linear distance plays no part in resolving ambiguities in Standard English might seem somewhat surprising, in view of the important part played by surface order in signaling the roles of various arguments. However, there is no mystery here: surface order is perceptually important in role-discovery, and in nothing else. Moreover, it is only important *to the extent that it contributes to role-discovery*. The latter point can be illustrated with an interesting observation of Lakoff's (1970c) to the effect that potential ambiguities are sometimes unacceptable, unless there is some feature with disambiguating effect. Thus he points out that ambiguity is not tolerated in (2.87b), although it is tolerated in (2.87a), and assumes that the distinction is due to the configurational peculiarities of the former. Moreover, these configurational factors do not lead to unacceptability in the presence of a disambiguating feature, like the fact that *Mary* is [-MALE] in (2.87c).

(2.87) a. John told Bill that he had won the Grand Prize.
b. *John and Bill walked in, and hé took off his coat.
c. John and Mary walked in, and hé took off his coat.

It seems that for ambiguity to be tolerated, the various readings must achieve a minimal level of 'distinctness.' Distinctness, in this sense, is an empirically discoverable notion, not one understood in advance, and undoubtedly deserves to serve the object of a separate study. It is sufficient, however, for our purposes to notice that the different surface ordering of the two putative antecedents of *he* assigns them different roles in (2.87a), while nothing of the kind is the case in (2.87b). Instead, surface order is irrelevant in a coordination, for, whatever the 'roles' of coordinate terms (if one accepts the Generative Semantics claim that logical *and* and *or* are predicates), these are clearly identical; moreover, coordinate terms are commutative. It would seem, therefore, that role-distinctness is required[24] in order for ambiguity to be allowed in such instances, and the different surface order of John and Bill in (2.87b) cannot provide the kind of distinctness required for ambiguity, precisely because it is not semantically or role-wise significant. This supports my earlier claim that linear distance in itself is irrelevant.

---

[24] In the absence of disambiguating information, role-distinctness between the two putative antecedents is a necessary, but not a sufficient, condition for ambiguity to be possible. An exhaustive study of this sufficiency condition is beyond the scope of this thesis, but the paradigm below suggests that (2.87b) is unacceptable for at least two reasons, namely (a) that the path connecting the antecedents goes through a coordinate node, and (b) that there is some node which dominates both antecedents and which is connected with the pronoun by a path which traverses a coordinate node.

(i) *John and Bill decided that he must go.
(ii) *John saw Bill lying on a sofa, and he dropped dead soon after.
(iii) *That Bill loves Mary and that Joan loves Jack seems odd to him.

I believe that the solution to the unacceptability of (2.87b) should be sought by investigating the notion of 'distinctness' I proposed above, and not by imposing a transderivational constraint, as Lakoff proposes. Lakoff's solution is, I think, inadequate, because the problem is not statable in purely grammatical terms. Thus, the transderivational constraint should be formulated in such a way as to allow all cases where there is a disambiguating factor, as in (2.87c), and this factor need not be strictly linguistic. For example, (2.88a) and (2.88b) seem to me perfectly acceptable, although the disambiguation is effected by extra-linguistic information; this information is context-supplied in the former case, and presupposed by the speaker in the latter case.

(2.88) a. John and Bill walked into and out of the room where I was sleeping respectively, and then hé struck me a big blow on the head.

b. Mao-Tse-Tung and Jack had a nice chat together, and hé started boasting he had made China the most powerful nation on earth.

I do not see how a grammar could have access to the disambiguating information in cases like (2.88). In (2.88b), for example, the grammar would have to be able to check whether what is asserted of the pronoun is true of both putative antecedents, and this would require the ascription to Mao-Tse-Tung of a feature like [+BOASTS

(iv) John$_i$ spoke to Bill$_j$ and to $\left\{\begin{matrix}\text{his}_i\\ \text{his}_j\end{matrix}\right\}$ mother.

(v) That John$_i$ and $\left\{\begin{matrix}\text{his}_i\\ \text{his}_j\end{matrix}\right\}$ mother were sick surprised Bill$_j$.

(vi) Bill$_i$ is sick, but neither John$_j$ nor $\left\{\begin{matrix}\text{his}_i\\ \text{his}_j\end{matrix}\right\}$ wife are.

Whatever the explanation for conditions (a) and (b) above, (b) is the same condition which prevents a discourse-boundary (like full stop, question mark, etc.) to intervene between the two antecedents on the one hand and the pronoun on the other, although such a boundary is allowed between the antecedents, as shown by (vii) and (viii) respectively.

(vii) a. Has John spoken to Bill about the impending bankruptcy?

b. *Yes, and his mother was most upset.

(viii) a. Does John$_i$ know about the bankruptcy?

b. Yes, and Bill$_j$ has told $\left\{\begin{matrix}\text{his}_i\\ \text{his}_j\end{matrix}\right\}$ mother too.

HE HAS MADE CHINA THE MOST POWERFUL NATION ON EARTH]!

2.2.1.3. In section 1.3.3.1, I mentioned Chomsky's claim that 'condition (1.102.i)' (which I reproduce below for convenience) 'in some cases, has the effect of reducing ambiguity, or to put it differently, of increasing the reliability of a reasonable performance strategy which seeks the nearest NP to a verb (or the head noun of a nominal phrase) as its subject.'

(1.102) i. No rule can involve *X*, *Y* in
...X...[...Z...-*WYV*...]...
where (i) *Z* is the subject of *WYV* and is
not controlled by a category containing *X*.

The cases of possible ambiguity which Chomsky has in mind involve a transformation proposed in Dougherty (1970) which derives (2.89a) from (2.89c) through (2.89b).

(2.89) a. The men hated each other.
b. The men each hated the other(s).
c. Each of the men hated the other(s).

Chomsky claims that *each*-movement must sometimes operate across sentence-boundaries, specifically in those cases where the subject of the embedded clause is a PRO-form controlled by a category in the matrix clause. Thus, he requires that (2.90a) be derived from (2.90b) and not from (2.90c), on grounds of synonymy.

(2.90) a. We want to kill each other.
b. Each of us wanted to kill the other(s).
c. We wanted each of us to kill the others.

Condition (1.102.i) serves to prevent the derivation of the b-sentence from the a-sentences in (2.91), where *Z* is not controlled by anything, being non-specific, in (2.92), where *Z* is controlled by the wrong category, and in (2.93), where *Z* is a specified non-controlled NP.

(2.91) a. Each of us heard about plans to kill the other(s).
b. *We heard about plans to kill each other.

(2.92) a. Each of us ordered John to kill the other(s).
b. *We ordered John to kill each other.

(2.93) a. Each of us expected the soldier(s) to shoot
the other(s).
b. *We expected the soldier to shoot each other.
c. We expected the soldiers to shoot each other.

(2.93c) is one of the cases Chomsky claims would be ambiguous without (1.102.i), since it could then be interpreted as derived either from (2.93a) or from *we expected each of the soldiers to shoot the other(s)*. In that case, one might claim that there is a minimal distance strategy which, at least in part, caused the introduction of (1.102.i) into the grammar. This would contradict my earlier proposal that minimal distance is not an operative criterion in English. However, I see some rather serious problems with such a proposal.

First, one would have to explain why the strategy was grammatized not just in (2.93c), but also in (2.93b), where no risk of ambiguity exists. Rule-generalization arguments seem to me rather forced in this instance.

Second, as Chomsky himself notes in his note 31, the input and output to the *each*-movement rule are not always synonymous, since *they shot each other* does not imply *each of them shot the other(s)*, although *they hate each other* does imply *each of them hates the other(s)*. Similarly, it seems to me that *we promised to respect each other* is consistent with a situation in which one person performed an act of promise on behalf of a group, while *we each promised to respect the other(s)* implies that each member of the group promised individually. The lack of synonymy in some cases does not argue against the *each*-movement rule, since transformations are allowed to change meaning in Chomsky's Extended Standard Theory, but it does weaken the motivation for a minimal distance strategy, by reducing the number of possible ambiguities.

Third, I find the rule of *each*-movement itself doubtful, as I cannot think what the source of sentences like those in (2.94) would be, if *each* must originate as a quantifier of some NP to the left of its surface position.

(2.94) a. John talked to Bill about killing each other.
b. Bob showed Jill a picture of each other.

It seems to me that both the *each*-movement and condition (1.102.i) can be dispensed with, and that all we need to do is to allow the generation of *each other* in its surface position, with the requirement that the expressed or understood subject of the clause *each other* is it be a plural or a collective. This would automatically explain the ungrammaticality of (2.91b), since unspecified subjects are not necessarily understood as plurals; in fact, if *Unspecified Subject* is understood as *someone*, it will be a singular.

2.2.2. A more sophisticated minimal distance principle, based not on linear distance, but on path-length measured along tree-branches, was proposed in Rosenbaum (1967) in order to determine the unique antecedent of a deleted complement subject. This principle was, however, convincingly argued not to exist in Postal (1968b), where it was shown that deleted complement subjects have no unique antecedents in general, as shown in (2.95a), and that antecedent-uniqueness was determined by independent factors in the subset of cases where it was found.[25] Moreover, even when the antecedent is unique, the minimal distance principle does not always correctly predict it, as demonstrated by the contrast between (2.95b) and (2.95c)--a minimal pair discovered by Jespersen.

---

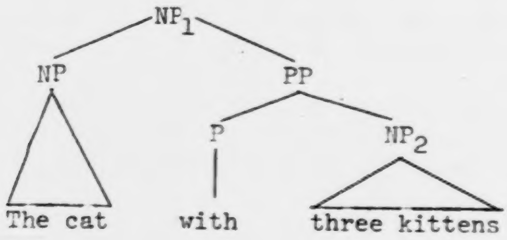
[25] A natural characterization of this subset was attempted in Grosu (1970).

(2.95) a. John talked to Bill about kissing Bertha.
b. He allowed her to go.
c. He promised her to go.

2.2.3. In sections 2.2.1 and 2.2.2 I argued against a strategy making use of linear distance or of distance measured along tree branches. Other minimal distance principles could be devised, and in this section, I shall inquire whether there is a principle that resolves ambiguities in terms of node-height. Consider, for example, (2.96), where the underlined constituent has the structure exhibited in (2.97).

(2.96) This is the cat with three kittens, which caught a lot of mice last week.

(2.97)



The non-restrictive relative clause could modify either three kittens or the cat with three kittens. In neither case would linear distance make a difference, for the relative clause is adjacent to both putative heads. Neither would branch-counting prefer one reading, for the same number of branches would connect the relative clause to its head, regardless of whether the latter is NP_1 or NP_2. However, NP_1 is higher up on the tree, and one may conceive of a strategy associating the relative clause with the lower node NP_2. As far as I have been able to ascertain, (2.96) is ambiguous and offers no support for such a strategy.

2.2.4.1. I shall now consider a number of cases in which linear distance appears to effect disambiguation. Perlmutter, at the 1970 Conference on English Syntax, proposed a transderivational constraint to handle the unambiguity of cases like (2.98a), which most people cannot perceive as a paraphrase of (2.98b).

(2.98) a. A woman hit a girl who was pregnant.
b. A woman who was pregnant hit a girl.
c. A woman just left who was pregnant.
d. A woman hit the curb who was pregnant.

Thus, although one would expect the relative clause in (2.98a) to be perceived as having arisen in one of two different ways, i.e., by being generated as a modifier of a girl or by Extraposition-from-NP, the latter reading seems to be out. An examination of the entire paradigm in (2.98a) will quickly reveal that the reading involving Extraposition-from-NP is out just in case ambiguity could arise. Perlmutter's transderivational constraint was formulated roughly as follows:

> (2.99) A derivation D is ill formed if Extraposition-from-NP applies in D, and if, by virtue of this, the last line of D becomes stringwise identical with the last line of D', where D and D' differ in semantic representation.

I do not believe that a transderivational constraint is the correct solution in this case, because there are dialects (e.g., A. Zwicky's) in which (2.98a) is ambiguous, but in which the reading ascribing pregnancy to the girl is far more likely. The greater likelihood of one reading cannot be captured in the transderivational constraint, which can either allow or disallow a reading. I believe that, as in other cases we have already seen, there is a perceptual principle at work (which is grammatized in Perlmutter's idiolect but not in Zwicky's), namely, (2.22), according to which simpler readings are preferred to more complex ones. The greater complexity of the structure resulting by Extraposition-from-NP in (2.98a) is due not to minimal distance considerations, but to strategy (2.37). As non-discontinuity is the extreme case of non-complexity of the intervening material, (2.37) correctly predicts that one reading is far less complex than the other.

There is, however, a far more serious problem with (2.99), namely, that it predicts that (2.100) is ill-formed on both potential readings.

(2.100) A woman sent a girl spinning who was pregnant.

This is so, because on either reading (2.100) arises through Extraposition-from-NP, and thereby becomes identical with the surface structure of the other potential reading. In contrast, (2.37) correctly predicts that the reading associating the extraposed relative with the girl is far more likely than the other (or the only possible one, if grammatization has occurred).

2.2.4.2. (2.37) predicts that, all other things being equal, a continuous reading will be preferred to one involving discontinuities whenever this situation arises, not only when it is brought about by Extraposition-from-NP. Another rule which can create this kind of situation is Particle-Movement. (2.101a) shows that the resulting discontinuity is possible, with some awkwardness when the intervening material is fairly complex, but that it is extremely unlikely (probably out in most dialects) when a continuous reading is possible, with the result that (2.101b) cannot in general be read as a paraphrase of (2.101c).

(2.101) a. ?John pushed the little girl who arrived yesterday down.
b. John pushed the little girl who had fallen down.
c. John pushed down the little girl who had fallen.

It is interesting that the discontinuous reading can be forced in (2.101b) by pausing between fallen and down, since the continuous reading, although less complex, becomes thereby ungrammatical. This cannot be done, however, in (2.98a), since a pause before the relative clause could be interpreted as marking the latter as a non-restrictive one, and the continuous reading would be possible and preferred.

2.3. I have attempted in section 2.1 to defend the claim that the complexities of erroneous closure, perceptual conflict, and interrupted behavior have linguistic relevance. Such perceptual limitations will be argued in Chapter Four to be significantly involved in Ross' constraints (as reanalyzed in Chapter Three).

The most inportant conclusion of section 2.2 for what follows is that certain remote relations (such as those between discontinuous components) are subject to perceptual limitations, while others (such as antecedent-anaphor relations) are not. This distinction will turn out to be crucial in proposing an explanation for the Dichotomous Behavior Principle.

CHAPTER III

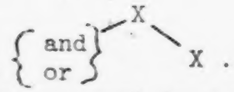
THE 'NUCLEUS-AND-SATTELITE' CONSTRUCTION

3.0. The two main purposes of this chapter are to argue that complex noun phrases, symmetric and asymmetric coordinate structures, and sentences or verb phrases modified by optional adverbials are special instances of a construction type which I call the 'Nucleus-and-Satellite' construction, and to show that they obey two constraints weaker than Ross' Island Constraint (see 1.4) (which predicts the complete freezing of 'A-immediately-over-A' structures).

3.1. The Nucleus-and-Satellite construction is defined in (3.1), and the constraints to which it is subject are stated in (3.2).

(3.1) Given three nodes X, Y, Z, such that
  (i) X and Y belong to the same grammatical category,
  (ii) X is properly analyzable as Y, Z,[1] and X immediately dominates both Y and Z,
  (iii) Z is optional, both in underlying and surface structure,
We call X a Nucleus and Satellite construction (N & S), Y a Nucleus, Z a Satellite

(3.2) Definition: Given a transformation T and a constituent C,
  a. If T moves C, T chops C, and
  b. If T moves an element of C, T maims C.
With respect to movement transformations, N & S's exhibit the following paradigm:
  (i) The Nucleus cannot be chopped
  (ii) The Nucleus can be maimed
  (iii) The Satellite can be chopped
  (iv) The Satellite cannot be mained
from under the domination of the N & S node.

(3.1i) is the A-over-A condition, while (3.1ii) includes the condition of immediate domination which Ross introduced in his Island Constraint. (3.1iii) is the condition I have added in order to keep out prepositional phrases, usually analyzed as $[^{P}_{NP}$ NP], complement sentences, often represented as $[^{Czer}_{S}$ S], and second coordinate terms, on the assumption that they are analyzed as  . The reasons for this added condition will become apparent in a moment.

---

[1] The linear order of Y and Z is not relevant

(3.1) is an effort to capture the notion of *Head and optional Dependent*. Both the notion of Head and that of optionality of the Dependent are hard to define operationally, although there have been numerous attempts in that direction (a succint summary of several structuralist and transformationalist attempts is given in J. Robinson's "Dependency Structures and Transformational Rules", in which she proposes a dependency-grammar solution). Although there is a string intuitive feeling that the head of the construction *the dog* is *dog* rather than *the*, it is not at all obvious how such a claim could be justified. Robinson proposes that the head is always obligatory, while the dependent may or may not be, and although this test is as good as any other I am aware of, it is by no means a perfect one. The difficulty is that it cannot be applied by simply inspecting a given structure, but requires an examination of all possible similar structures; some of the obvious difficulties are that we can never be sure we have inspected *all* the pertinent structures, we do not have a procedure by which to identify 'corresponding' elements in similar structures, etc. In fact, Robinson regards the head as the 'characterizing' element of a construction, which is undoubtedly an intuitively satisfactory definition, but not one that makes possible a mechanical decision. As I do not have a mechanical procedure either, I shall settle for the hope that intuitions as to which element is the head of a particular construction will in general agree.[2]

---

[2] Since Robinson's main interest was not to explicate the semantic feeling that some element is central to a construction, but rather to find a characterizing, necessary terminal element which could be identified as the 'governor' (in the sense of G. Lakoff (1970a)) of various transformations, she posits heads for *all* constructions, including those that have usually been regarded as exocentric. Thus, she has an abstract element T (sentence-Type) which constitutes the head of sentences, and she regards the preposition as the head of a prepositional phrase (because it is terminal). It is not entirely clear to me how the doubly underscored string in *that John will leave his wife is perfectly possible* would be analyzed in this approach; indeed, it is not only an S, but also a NP, and what the nominal head could be is not easy to see (although some abstraction like 'the event' or 'the proposition' could certainly be posited). In any event, this approach is not of interest to our present discussion. Although I have no non-intuitive arguments, I do not think it would be easy to make informants agree that a preposition is 'modified' by its NP, that a complementizer is 'modified' by the following sentence, or that a sentence 'modified' its Type. There is no conflict involved here, for Robinson makes it clear that she is talking about

'characterizing elements', not about 'centers'. I wish, however, to keep the notion 'head' for endocentric constructions only.

The optionality of constituents is also a notion definable 'in general' and with some appeal to intuition, but unfortunately not in a mechanical way either. Thus, it is not difficult to assent to the proposition that any term of a symmetric coordination is, in some sense, optional, since the other terms can, in principle, play the same role within the containing framework as the entire coordination does. The best operational test for optionality would be interchangeability in all contexts, but it is already a commonplace that this strong form of optionality is hardly, if ever, encountered. Thus, although a coordinate term and the entire coordination are often interchangeable, as in (3.3) and (3.4), counterexamples can easily be found, as in (3.5)-(3.7).

(3.3) a. That Bill murdered John and Mary is awful.
b. That Bill murdered Mary is awful.

(3.4) a. Either Bill or Mary will have to leave.
b. Bill will have to leave.

(3.5) a. John and Mary went their separate ways.
b. *John went his separate way.

(3.6) a. John, Mary and Bill are my three best friends.
b. *John and Bill are my three best friends.

(3.7) a. Bill, Sally and the merchant divided the money among themselves.
b. *Bill and the merchant divided the money among themselves.

Moreover, interchangeability in context would be a necessary, but not a sufficient, condition for a mechanical decision; thus, although the singly- and the doubly-underscored material in (3.8) are interchangeable with preservation of grammaticality, it seems counterintuitive to regard the singly-underscored material minus the doubly underscored one as optional.

(3.8) The man who talked to the lady with a friendly dog in her arms is coming towards you.

It appears that both the notion of head and that of optionality of constituents are theoretically ill-defined at present, although both are intuitively clear in a large number of cases. I shall use both these notions in what follows, because I believe that interesting generalizations hinge on them, without however pursuing the task of defining them formally any further. I have coined the terms Nucleus and Satellite because, as I shall argue below, coordinate constructions satisfy both the structural definition in (3.1) and the behavioral prediction in (3.2), and calling a coordination a head-and-dependent construction might appear as a more startling claim than it really is.

The behavior of N & S's with respect to movement transformations as predicted by (3.2) is clearly less rigid than the prediction made by the IC. Indeed, the latter predicts that neither Nuclei nor Satellites are either choppable or maimable, while (3.2) predicts that Nuclei can be maimed but not chopped, while Satellites can be chopped but not maimed.[3]

It can be seen already why condition (3.1iii) was imposed; indeed, prepositional phrases, complement sentences, second coordinate terms, and in general, constituents introduced by obligatory connectors, do not behave as predicted by (3.2). Indeed, the NP of a prepositional phrase can be freely moved in English, thus failing to comply with (3.2i), and the connector can never be moved, while Satellites in general can (as predicted by (3.2iii)). I suspect that the freezing of such connectors is a very strong, probably universal, constraint; although my knowledge of languages is quite limited, I would be very surprised if a rule moving connectors (whether these precede or follow the elements they introduce) turned out to exist in some language.

The problem of proposing an explanatorily adequate account of the conjunction of (3.1) and (3.2) will be taken up in Chapter Four, but it can be shown rather easily at this point that my proposals escape the objections I raised against the IC in (1.118). I show this in (3.9) below:

(3.9) a. In simple Extraposition, the extraposed S is maimable, although immediately dominated by the category S, because (3.1iii) is not satisfied.
b. Extraposition from NP is permitted by (3.2iii).
c. Extraposition-of-PP is permitted by (3.2iii).
d. Prepositions can be stranded because (3.1iii) is not fulfilled.
e. An NP complementing another NP can be chopped, because of (3.2iii).
f. Complement sentences can be maimed because (3.1iii) is not fulfilled.
g. The maiming of matrix S's modified by adverbial clauses is permitted by (3.2ii).
h. The chopping of adverbials is allowed by (3.2iii).

---

[3] (3.2ii) and (3.2iii) do not, however, make the claim that each language will have rules capable of maiming every single instance of Nucleus and/or of chopping every single instance of Satellite; the claim is merely that there is no general constraint blocking all movements out of N & S's, as the IC predicts.

(3.9) i. Adverb Preposing is allowed by (3.2iii).
j. The behavior of assymetric coordination will be shown to be entirely consistent with the conjunction (3.1), (3.2).
k. A complete answer is not possible at this point, but it will later be shown that across-the-board chopping is not an instance in which a Satellite is maimed 'from under the domination of the N & S node'.

The IC has thus been replaced by two weaker constraints, (3.2i) and (3.2iv). I shall now show that these two constraints are in fact largely independent, both with respect to the structures, and to the kinds of phenomena, to which they are applicable.

As far as structural applicability is concerned, (3.2iv) alone is restricted to Satellites, while (3.2i) is applicable to heads in general. Thus, if we accept the currently held views (which, as I said, I cannot prove) that the noun is the head of the NP and the verb is the head of the VP, there are hardly any rules which move just the constituent N or V. The fact that NP's can move out of PP's leaving a stranded preposition is not a counterexample, since, as I pointed out before, I consider the notion 'head' applicable to endocendric constructions only (this is precisely the intuition which has led previous writers to posit the endocentric/excentric dichotomy), unlike Robinson, who posits heads in exocentric constructions like PP as well. On the other hand, (3.2iv) cannot be extended to modifiers in general, for, if VP is regarded as an endocentric construction with the V as head, a complement of V can most certainly be maimed.

With regard to the sets of phenomena which obey (3.2i) and (3.2iv), neither of which is restricted to chopping, it is interesting that they do not coincide. The phenomena which obey (3.2iv) have been discussed in Chapter One, as they constitute one member of the partition determined by the Dichotomous Behavior Principle. Which phenomena obey (3.2i) has not been studied in detail, but it is sufficient to show that one kind of phenomena obey one constraint and disobey the other in order to show that these constraints determine different partitions of syntactic processes. (3.10a) and (3.10b) show that any neutralization 'into' complex NP's is constrained by (3.2iv), but not by (3.2i).

(3.10) a. *I never heard of the claim that John ever killed anybody.
b. I never knew anybody who claimed that Bill was unfriendly.

There may well be other rules which behave differently with respect to the two constraints, but the facts are not entirely clear to me. Thus, it would appear that copying rules, which we recall are exempt from (3.2iv), are subject to (3.2i), We exemplify this property in (3.11) with the rule of Relativization in Hebrew.

(3.11) a. Raita et haieled (,) she kvar holex lebeit-sefer.
'You saw the child (,) who already goes to school.
b. Haieled she raita oto (*(,) she kvar holex lebeit-sefer (,)) xaxam.
'The child who you saw him (*(,) who already goes to school (,)) is clever.

Referential pronouns, of the kind that result from copying, cannot in general be heads of relative clauses. The point is illustrated with respect to Hebrew and English in (3.12a) and (3.12b) respectively (the latter is a translation of the former).

(3.12) a. Raiti et haialda$_i$, veata gam raita ota$_i$ (*(,) she kvar holexet lebeit-sefer).
b. I saw the girl$_i$, and you also saw her$_i$ (*(,) who already goes to school).

It is not entirely clear if (and/or to what extent) (3.2i) is responsible for the unacceptability of (3.12). On the restrictive reading of the relative clause, the head does not refer independently of the N & S which it heads, and thus cannot be coreferential with anything. But it is not ovious why the non-restrictive reading of the relative clause should be bad (at least when the latter is a comment of the speaker, since some speakers accept (3.12b) with the relative expressing an after-thought; however, on the after-thought interpretation, there is no way of knowing whether the clause modifies the girl or her!). A possible reason for the unacceptability of (3.12a) and (3.12b) may be that it is odd to make a comment on an NP whose referent is not identified uniquely, and it seems to make more sense to attach a comment to the more fully specified antecedent, which has a better chance of achieving identification of the pertinent referent. This assumption is supported by the fact that (3.12c) is bad in general, but good if spoken with heavy stress on he and in the presence of the referent of he.

(3.12) c. He, who is my best friend, is very sick.

In the latter case, he is in fact a deictic, and deictics are perfectly all right with non-restrictive relatives; notice that deictics identify uniquely (I will argue below that precisely for that reason deictics cannot take restrictive relatives), and thus support my point. All this suggests that the badness of (3.12a) and (3.12b) may not be due to (3.2i) on either the restrictive or non-restrictive reading. It should also be noticed that reflexives function similarly:

(3.12) d. John cut himself (*(,) who is my best friend)).

(3.12d) is tolerable with the relative as an after-thought, but as in the case of (3.12b), it cannot be shown that the after-thought necessarily modifies the reflexive rather than the antecedent.

One additional reason for doubting that pronominalization is subject to (3.2i) is that non-referential pronouns can be heads of restrictive relatives, as (3.13) shows.

(3.13) I saw a pretty girl, and you saw one who wasn't too attractive.

The fact that (3.13) is acceptable ties in nicely with my hypothesis that referentiality was responsible for the unacceptability of (3.12b) on the restrictive reading.[4] It should also be noticed that (3.13) is bad if one is unstressed and the relative is non-restrictive; this also fits my hypothesis concerning the unacceptability of (3.12) on the non-restrictive reading, since unstressed one is non-specific, and therefore does not identify a referent. On the other hand, (3.13) is possible with a non-restrictive reading if one is stressed, for one is then specific and does identify a referent.

NP's modified by complement sentences behave essentially in the same way as NP's modified by relative clauses, that is, they cannot be referential pronouns, but can be non-referential ones, as (3.14) shows.

(3.14) a. *I heard the hypothesis that Bill has cancer and you heard it that Mary has pneumonia.
b. I heard a prognosis that Bill has cancer and you heard one that Mary has pneumonia.[5]
c. *I am aware of the fact that Bill is sick, but you are not aware of one that Mary is sick as well.

The ungrammaticality of (3.14c) is due to independent reasons; indeed, one is indefinite, and the structure *you are aware of a fact that Mary is sick is also ill-formed. In general, pronominalization with one in such cases is bad just in those cases in which the corresponding full NP is also bad (the constraints on the determiners of NP's with complement sentences are discussed at the end of section 3.2.3).

---

[4] Pronouns having the surface shape of referential pronouns like the one in (3.12b) can be used nonreferentially, in which case restrictive relatives are perfectly possible, as in (i):

(i) He who commits murder shall be severely punished.

[5] (3.14b) is good only if one is stressed (and, therefore, specific). Thus, complement sentences function like non-restrictive relatives, rather than like restrictive ones, which is consistent with the observation that complement sentences supply additional information on the head NP, but do not provide narrower identification.

Concerning copying rules, I wish to point out that if they are made subject to (3.2i), it becomes possible to explain a baffling fact about the CSC in languages with copying rules. Thus, an element of a coordinate term can be copied, but an entire coordinate term cannot. The situation is illustrated in (3.15a) and (3.15b), with examples from Hebrew (R. Wojcik informs me that the same situation is found in Breton).

(3.15) a. Haieled$_i$ she raita et ima shelo$_i$ ve et aba shel xaim iatsa.
'The child$_i$ who you say his$_i$ mother and Chaim's father has left.
b. *Haieled$_i$ she raita oto$_i$ ve et xaim iatsa.
'The child$_i$ who you saw him$_i$ and Chaim has left.

If, as I shall argue later, coordinate structures are N & S's, the paradigm in (3.15a) and (3.15b) is accounted for by (3.2iv) and (3.2i) respectively.

The behavior of copied pronouns differs from that of ordinary referential pronouns, as the perfect acceptability of (3.16), in contrast to the unacceptability of (3.15b), shows.

(3.16) Haieled babait, ve raiti et xaim ve oto korim sefarim.
'The child is at home, and I saw Chaim and him reading books.

The difference between ordinary pronouns and the ones that result from copying is that the latter are subject to coreferentiality neutralization. Thus, when such a situation arises in English (even by invited inference, as argued in Chapter One) the result is far from good, although less bad than when the pronoun is removed completely, as shown in (3.17).

(3.17) a. ?*The rock is too heavy for me to try to

$$\text{pick}\left\{\begin{matrix}\text{it and the hammer}\\\text{the hammer and it}\end{matrix}\right\}\text{up.}$$

b. **The rock is too heavy for me to try to

$$\text{pick}\left\{\begin{matrix}\text{and the hammer}\\\text{the hammer and}\end{matrix}\right\}\text{up.}$$

Although it is hardly possible to compare degrees of acceptability across languages, there is indirect evidence that structures like (3.15b) in Hebrew are less unacceptable than structures in which a coordinate term has been removed; in other words, the unacceptability of (3.15b) is like (3.17a), rather than like (3.17b). The reason for believing this is that structures like (3.15b) become acceptable under certain conditions, one of which is topicalization, as shown in (3.18).

(3.18) Haieled she oto ve et xaim her ?eti leha iatsa
min habait.
'The child that him and Chaim I showed you has
left the house'.

Although I am at a loss to explain why topicalization should improve acceptability in such cases, it is reasonable to assume that copying a head is less bad than chopping it. This is confirmed by evidence internal to Hebrew; question formation is a chopping rule, and we see that (3.19) is out, despite topicalization.

(3.19) *Mi ata xoshev she ve et xaim her?eti le imi?
'Who do you think that and Chaim I showed
my mother?'

The contrast in acceptability between (3.15b) and (3.16) reduces the likelihood that the unacceptability of (3.12a) or (3.12b) have much to do with (3.2i). On the other hand, the fact that reflexives and ordinary definite pronouns are both bad as heads of relative clauses (as shown by (3.12d) and (3.12b) respectively) does not mean that Reflexivization is not subject to (3.2i), for (3.12d) could be out both by (3.2i) and by the semantic constraints on ordinary pronominalization discussed above. If Reflexivization is made subject to (3.2i) but not to (3.2iv), it becomes possible to explain the contrast in acceptability between (1.92a) and (1.94). The following pair suggests even more strongly that Reflexivization obeys (3.2i) rather than the much stronger CSC, as Ross thought.[6]

(3.20) a. *John hit $\left\{\begin{array}{l}\text{Mary}\\ \text{the book}\end{array}\right\}$ and himself.
b. John put a knife and a picture of himself
on the table.

With respect to the copying of sentences modified by adverbials, I have been unable to find examples in Hebrew in which a S is copied. An illustration would be a sentence like *That Bill is crazy, mother believes it, if it existed. Notice, however, that the pronouns that would occur from such copying would be different from both referential pronouns and non-referential ones like one, since they would involve identity-of-sense anaphora. Although no identity-of-sense copying is found in English, it can be shown that such phenomena are constrained by (3.2i) quite independently of copying:

---

[6] If ordinary pronominalization and feature-changing rules are not subject to (3.2i), it is rather startling that Reflexivization, which is both a pronominalization and a feature-changing rule, be subject to it; nevertheless, the facts seem to suggest it.

In Chapter Four, where I suggest a natural explanation for (3.2iv) and (3.2i), I argue that there is a unique principle which determines which phenomena are subject to the former. With respect to the latter, I have doubts that Reflexivization and chopping rules obey it for the same reason, as will become clearer in 4.5, where the explanation for chopping rules being subject to (3.2i) is offered.

(3.21) a. John resents Mary's going around naked, but Bill does not resent Mary's going around naked for purposes of publicity.
b. *John resents Mary's going around naked, but Bill does not resent it for purposes of publicity [where for purposes of publicity modifies it, not resent].

Notice that there is a form of identity-of-sense pronominalization, which is to one -pronominalization what it in (3.21b) is to referential pronouns, as (3.21c) shows. Presumably, it in identity-of-sense phenomena implies total identity (as referential pronouns do), unlike pro-forms like one and the underscored constituent in (3.21c).

(3.21) c. John resents Mary's going around naked, but Bill does not resent such behavior for purposes of publicity.

In conclusion, it is not clear to me to what extent the copying and pronominalization of Nuclei is constrained by (3.2i) and by independent (semantic) constraints on pronominalization. Whatever the situation, it is clear that (3.2i) and (3.2iv) differ with respect to the phenomena they affect, as was earlier shown with feature-changing rules.

In this section, we have discarded the notion of A-over-A configuration (defined by (3.1i) and (3.1ii)) in favor of the notion of N & S (defined by (3.1i), (3.1ii), and (3.1iii)) and of head (undefined but assumed to be understood intuitively). We have also weakened the prediction that the structures thus defined are totally impervious to certain syntactic operations by introducing the more limited constraints (3.2i) and (3.2iv).

The remainder of this chapter, that is, the sections 3.2, 3.3, and 3.4, is devoted to complex NP's, optional adverbials modifying S's or VP's, and coordinate structures respecitvely. Each section is broken down into subsections corresponding to the basic subtypes found for each construction type. In each subsection, it is argued that the construction under discussion satisfies the requirements in (3.1) (and therefore is a legitimate N & S), and that it behaves, with respect to movement transformations, as predicted by (3.2i)-(3.2iv).

3.2. In this section, which is devoted to complex NP's, I discuss non-restrictive relatives, restrictive relatives, and nominal complements, in that order.

3.2.1. We begin by considering non-restrictive relatives, as in (3.22).

(3.22) A girl, who you know, just left.

3.2.1.1. I shall first argue that the underlined string in (3.22) satisfied the first part of (3.1ii), that is, that it is properly analyzable into two constituents, for it is conceivable that it might have the structure DET N S. The tests for showing that the string DET N forms a constituent in this case are not numerous; indeed, the usual way of proving that a string is a constituent is by showing that it can be moved by various syntactic processes, but the constituent at issue is frozen by (3.2i). Tests can nevertheles be found; I offer three below.

The first test concerns the output of Coordination Reduction when this rule applies to the right. Indeed, when the output contains a coordination of non-constituents, the latter are obligatorily separated by pause, as we can see in the contrast between (3.23) and (3.24).

(3.23) a. John and Mary ate the cake.
b. John, and Mary, ate the cake.

(3.24) a. John baked, and Mary ate, the cake.
b. *John baked and Mary ate the cake.

Consider now (3.25). If the underscored strings in (3.25a) were not constituents, we would expect (3.25b) to be ungrammatical, like (3.24b). The fact that it is grammatical suggests that (3.25b), like (3.23a), contains a conjunction of constituents.

(3.25) a. The boy, who wants to go to Paris, and the girl, who wants to go to Paris, are in love with one another.
b. The boy and the girl, who want to go to Paris, are in love with one another.

The second test concerns a 'hierarchical constraint' on identity deletion in coordinate structures proposed in J. H. Y. Tai (1971):

(3.26) Higher identical constituents must be deleted before lower identical constituents.

On the assumption that the underscored strings in (3.25a) are constituents in construction with the following relative clauses, (3.26) can explain the grammaticality of (3.27a) and the ungrammaticality of (3.27b).

(3.27) a. The boy and girl, who want to go to Paris, are in love with one another.
b. *The boy, who wants to go to Paris, and girl, who wants to go to Paris, are in love with one another.

Indeed, (3.27a) can be derived from (3.25a) through the intermediate stage (3.25b), while (3.27b) is derived directly from (3.25a). The contrast in grammaticality can be accounted for by (3.26) if the conjuncts of (3.25a) are analyzed as $[^{NP}_{NP}\ S]$, but not if they are analyzed as $[^{DET}_{NP}\ N\ S]$ or $[^{DET}_{NP}\ [^{N}_{N}\ S]]$. The constraint on Coordination Reduction is apparently stronger than predicted by (3.26), for lower order constituents cannot be reduced to the left even if they are the only identical ones, as shown in (3.28). For reasons I do not understand, the reduction of lower constituents to the right exhibits a gradual decrease in acceptability, rather than abrupt ungrammaticality, as shown in (3.29).

(3.28) a. The boy, who wants to see Paris, and the girl, who intends to visit London, are in love with one another.
b. *The boy, who wants to see Paris, and girl, who intends to visit London, are in love with one another.

(3.29) a. Mary baked, and John ate, the cake.
b. ?I claim that Mary baked, and you think that John ate, the cake.
c.?*I claim that my sister believes that Mary baked, and you think that your mother imagines that John ate, the cake.

The third test is of more limited interest, since it shows that structures like (3.22) cannot have the structure $[^{DET}_{NP}\ [^{N}_{N}\ S]]$, but it does not show that they cannot have the structure $[^{DET}_{NP}\ N\ S]$. I shall present it nevertheless, since the elimination of even one alternative is worth considering.

The heads of genitives can be deleted under certain conditions which need not concern us here. Thus, (3.29a) can become (3.30b).

(3.30) a. I know Jill's boy-friend, but not Mary's boy-friend.
b. I know Jill's boyfriend, but not Mary's.

If constructions like (3.22) had the internal structure DET $[^{N}_{N}\ S]$, we would expect (3.31b) to be derivable from (3.31a), and (3.32b) not to be derivable from (3.32a). Both expectations are thwarted.

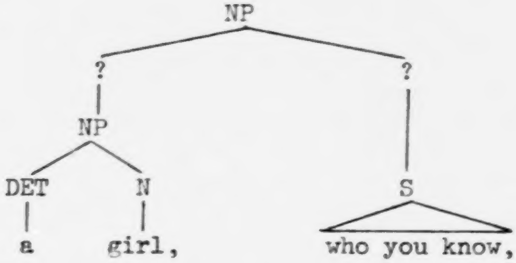
(3.31) a. I know Jill's boy-friend, who is from Columbus, and you know Mary's boy-friend, who is from Columbus.
≠ b. I know Jill's boy-friend, who is from Columbus, and you know Mary's.

(3.32) a. I know Mary's boy-friend, who is a doctor, and you know Mary's boy-friend, who is an engineer.

(3.32) b. I know Mary's boy-friend, who is a doctor,
and you know Mary's, who is an engineer.

Having shown that structures like (3.22) satisfy the first part of (3.1ii) we have in fact shown that they satisfy the second part of (3.1ii). True, the preceding arguments would be consistent with an analysis of (3.22) such as in (3.33), but we can avoid the latter by imposing the condition (3.34) which, as far as I am aware of, has been assumed by all writers in transformational grammar.

(3.33)



(3.34) Never posit more structure than you have evidence for.

With respect to (3.1i), notice that there is an implicational relation between it and (3.1iii). Thus, the latter implies the former, although the opposite is not true, as shown by prepositional phrases, sentences in construction with complementizers, etc. Therefore, if we can show that (3.1iii) holds, we no longer need to show that (3.1i) holds. I believe that non-restrictive relative clauses are among the clearest cases of optional constituents, although, as I pointed out earlier, I know of no operational tests for proving optionality. Thus, interchangeability in all contexts with preservation of grammaticality would be a good test (if appropriately constrained to avoid counterexamples like (3.8)), but it would be inapplicable in any event, since we could never look at *all contexts*; besides, the requirement could presumably be shown not to hold even for non-restrictive relatives, by cases like (3.35).

(3.35) a. Mary, who sleeps with anybody who asks her,
is planning a trip to Europe, and believe
it or not, her mother doesn't even know
that she is that kind of girl.
b.??Mary is planning a trip to Europe, and believe
it or not, her mother doesn't even know that
she is that kind of girl.

The relevance of optionality to the problem at issue is discussed at greater length in Appendix Two.

3.2.1.2. We now turn to the relation between structures like (3.22) and (3.2). (3.2i) is illustrated in (3.36).

(3.36) a. I told John, who is sick, that he cannot go on living here.
b. *It's John that I told, who is sick, that he cannot go on living here.

(3.2ii) appears not to be realized in English, since there are no rules which maim the configuration [DET N]. There is no violation involved, however, for (3.2i), as I have pointed out already, applies not just to Nuclei, but to Heads in general. Therefore, the constituent N is blocked by (3.2i); the constituent DET is blocked by the Left Branch Condition in English, but apparently not in Latin or Russian.

(3.2iii) is illustrated by Extraposition-from-NP, although the latter is not accepted with non-restrictive relatives in all dialects. Thus, (3.37) is acceptable to some speakers and unacceptable to others.

(3.37) a. A girl, who I couldn't care less about, has just left.
b. A girl has just left, who I couldn't care less about.

(3.2iv) is illustrated by (3.38).

(3.38) a. I know a girl, who is greatly interested in Bill.
b. *Who do I know a girl, who is greatly interested in?

3.2.2.1. The three tests involving Conjunction Reduction that we used in arguing that non-restrictive relative constructions satisfy (3.1ii) are mechanically applicable to restrictive constructions. The same examples (3.25), (3.27), (3.28) and (3.31)-(3.32) with the appropriate commas removed can be used to support the corresponding arguments.

With regard to requirements (3.1i) and (3.1iii), they are also satisfied in most instances. There are, however, a few apparently embarrassing cases, which I discuss below.

Thus, nouns like time, place, manner seem to require restrictive relatives when they are preceded by certain determiners, and particularly when they form optional adverbials. Relative clauses are apparently required by a, the, every, optional with some, and usually impossible with that. However, it would be rather unnatural to claim that the syntactic representation of these few cases is radically different from that of other restrictive relative constructions. First, notice that the optionality of the relative clause seems to depend upon the optionality of the higher order construction which it forms with its head. Thus, while the sentences in (3.39) are always bad, the sentence in (3.40) are acceptable, if some previous context is furnished.

(3.39) a. *He arrived at a time.
b. *He killed John at a place.
c. *He slapped Midge in a manner.

(3.40) a. He proposed a time.
b. He went to a place.
c. He decided on a manner.

The most interesting pair is (3.39b)-(3.40b), since both are adverbials distinguished only by optionality. Thus, the counterargument that one may raise to the effect that a time and a manner can be pronominalized with it in (3.40) but not in (3.39) is irrelevant here, since neither (3.39b) nor (3.40b) can be pronominalized with it. But if the pertinent sequences have the same structure in (3.39) and in (3.40), it can be shown that they are all NP's (i.e., that (3.1i) is satisfied), since they can be conjoined with other NP's as in (3.41).

(3.41) a. He mentioned the time and the person that I was interested in.
b. We discussed the woman and the place that I had found out about.
c. I told him about the principle and the manner that he was curious about.

I wish to suggest that the sentences in (3.39) are out not because they are semantically ill-formed, but because they are semantically odd, or, more specifically, because they violate an accepted conversational principle like (2.23), which I repeat below for convenience:

(2.23) When something is done for a purpose, and that purpose cannot be achieved in principle, perceptual conflict arises.

Indeed, the adverbials in (3.39) are optional in surface structure, since it can be inferred that when someone does something, he does that at some time, at some place, and in some manner. Therefore, if one goes to the trouble of mentioning the adverbials explicitly, the assumption is created that these will carry additional information, and this turns out not to be the case. Notice that the presence of completely uninformative optional constituents is highly unacceptable in other cases as well, in fact, to such an extent that it is difficult to tell whether the unacceptability of (3.42) has been grammatized or not (the offending constituents are underscored).

(3.42) a. ?*The car was given away to a human being.
b. ?*The president has been murdered by a human being.

Notice that the sentences in (3.39) and (3.42) remain bad if a is replaced by the in the underscored portions. The reason is basically the same, I think, since the is referential in non-opaque contexts like the above, but fails to identify uniquely if unqualified by a relative clause; therefore, the above sentences with the in the appropriate places fail to fulfill their purported purpose, and are out too, as predicted by (2.23).

The problem with every seems to be a little different. Thus, every seems to be bad not just in sentences like (3.39), but also in sentences like (3.40), as shown by (3.43) and (3.44) respectively. The reason is, I suggest, the implausibility of the assertion made.

(3.43) a. *He visits her at every time.
b. *He kissed Mary in every place.
c. *He tortured her in every manner.

(3.44) a. *He mentioned every time.
b. *He went to every place.
c. *He discussed every manner.

An apparently idiosyncratic fact for which I have no explanation is that all is not subject to the same restrictions as every. Thus, he visits her at all times is acceptable, unlike (4.43a), because the former can be understood with some qualification, such as "he visits her at all conceivable times", or "he does so at all times when he is available", etc.

The behavior of that is fairly clear. Thus, apart from certain constructions like he is looking for that which nobody has yet discovered, that is referential, and if so, it identifies uniquely. Therefore, the addition of a restrictive relative, whose purpose is precisely to achieve unique identification, appears as a superfluous endeavor and is ruled out by (2.23). That the unacceptability of referential that with restrictive relatives, as in (3.45), is not an idiosyncratic property of that but rather follows from some general principle like (2.23) is shown by the fact that restrictive relatives are unacceptable in two other cases in which that is not found, but uniqueness of reference is achieved independently of the relative clause. The first case involves the oddity of constructions consisting of a head NP and a restrictive clause which in turn form the head of construction of the same kind. The second case concerns the unacceptability of restrictive relatives modifying NP's whose referents are empirically known to be unique. These two unacceptable constructions are illustrated in (3.46) and (3.47) respectively.

(3.45) *I showed John that house which is white.

(3.46) *[[This is the boy who is tall] who you saw
yesterday].

(3.47) *I showed Midge the sun which burns fiercely.

(3.45) is unacceptable because both the deictic properties of that and the relative clause coupled with a definite head NP can identify uniquely, and the fact that both devices have been used suggests that one of them cannot identify uniquely, which conflicts with the presupposition of uniqueness associated with either. (3.46) is bad because the second relative clause also suggests that its NP head can be more narrowly specified, and this again contradicts the presupposition of the first relative clause. Finally, in (3.47), even though there is no linguistic feature about the sun requiring that it have a unique referent, it is an empirical assumption that there is only one sun, and the restrictive relative is just as bad as in (3.45) or (3.46).

It remains to explain the behavior of some, since it is not obvious why the sentences in (3.39) should be out, while those in (3.48) should be acceptable.

(3.48) a. He arrived at some time.
b. He slayed Mary at some place.
c. He stole the money in some manner.

To begin with, notice that the sentences in (3.48) are acceptable only if some is stressed. If some is unstressed, a sentence like (3.48a) can only be used as an evaded answer to when did he arrive? If some is stressed, however, sentences like those in (3.48) can be used when the speaker does not wish to reveal certain facts to the hearer, or when he does not wish the facts in question to be overheard by a third party and, perhaps, hopes that the hearer can guess what he has in mind, etc. The difference in acceptability between (3.39) and (3.48) is that some, unlike a, constitutes an explicit signal to the effect that some information will not be revealed, and therefore (3.48) is not ruled out by (2.23), since there is no longer any reason to believe that the adverbial is used for the purpose of adding information.

3.2.2.2. The applicability of (3.2i) is illustrated by (3.49):

(3.49) a. I showed the boy who was from Chicago to the girl with a red hat.
b. *It's the boy I showed who was from Chicago to the girl with a red hat.

(3.2ii) is inapplicable in English, but this is not a counterexample; the reasons are the same as those given in the discussion of non-restricted relatives.

(3.2iii) is illustrated by Extraposition-from-NP, as shown in (3.50):

(3.50) a. A man who was drunk walked in.
b. A man walked in who was drunk.

The applicability of (3.2iv) is shown in (3.51);

(3.51) a. I showed you a boy who likes bananas.
b. *What did I show you a boy who likes?

3.2.3.1. While my earlier claim that relative clauses with heads have the surface structure representation $[_{NP}^{NP}\ S]$ is, as far as I know, the view held by most writers in transformational grammar, it has been proposed by a number of authors (Chomsky (1970), (1971) in particular) that nouns with complement sentences have the surface structure representation $[_{NP}^{DET}\ [_{N}^{N}\ S]]$. While the motivations for positing such a structure are fairly clear (the desire to express the parellism between *John's claim that Mary is dead* and *John claimed that Mary was dead*), there is clear syntactic evidence against positing the surface structure that Chomsky defends.

To begin with, the three tests that I used in arguing that relative clauses are properly analyzable into two constituents yield the same results for nouns with complement sentences.

First, (3.52) exactly parallels (3.25), as (3.52b) is grammatical even though there is no pause between the underscored conjuncts (as we recall, non-constituents cannot be coordinated without intervening pause).

(3.52) a. The suspicion that John may be a spy and the claim that John may be a spy were both discussed by the FBI investigators.
b. *The suspicion and the claim* that John may be a spy were both discussed by the FBI investigators.

Contrast also (3.53a) and (3.53b):

(3.53) a. John believes, and Mary claims, that Ike is stupid.
b. John's belief and Mary's claim that Ike is stupid are odd.

Notice that the second conjunct *must* be flanked by pauses in (3.53a) but *not* in (3.53b), which suggests that only the conjuncts in the former are non-constituents.

Second, if the determiner were indeed "the subject" of the NP, as Chomsky claims, then we would expect (3.54b) to be grammatical. This is so, because (3.45a), a clear case of determiners in construction with the constituent N, does indeed reduce to (3.55b). Moreover, the grammaticality of (3.56b) in conjunction with the ungrammaticality of (3.54b) argues against Chomsky's claim that NP's and S's have the same internal structure except for node-labels.

(3.54) a. $\left\{\begin{matrix}\text{John's}\\ \text{The}\end{matrix}\right\}$ claim that Bill was a spy and $\left\{\begin{matrix}\text{John's}\\ \text{the}\end{matrix}\right\}$ insinuation that Mary had recruited him came as a big surprise.
b. * $\left\{\begin{matrix}\text{John's}\\ \text{The}\end{matrix}\right\}$ claim that Bill was a spy and insinuation that Mary had recruited him came as a big surprise.

(3.55) a. John's ideas and John's beliefs are shocking.
b. John's ideas and beliefs are shocking.

(3.56) a. John claimed that the boys were dead and John insinuated that Mary had murdered them.
b. John claimed that the boys were dead and insinuated that Mary had murdered them.

Third, recall the rule which deletes the constituent N in construction with a genitive determiner. If Chomsky's analysis were correct, we would expect (3.58a) to reduce to (3.58b) in the same way in which (3.57a) reduces to (3.57b). However, (3.58a) and (3.58b) are not paraphrases, since the missing constituent in the latter can only be understood (if at all) as *claim* but not as *claim that Mary deceived Bill*.

(3.57) a. I like John's ideas, but not Mary's ideas.
b. I like John's ideas, but not Mary's.

(3.58) a. I heard of John's claim that Mary deceived Bill, but not of Jack's claim that Mary deceived Bill.
b.??I heard of John's claim that Mary deceived Bill, but not of Jack's.

Moreover, if the construction at issue had indeed the structure $[_{NP}\ \mathrm{DET}\ [_{N}\ \mathrm{N}\ \mathrm{S}]]$, that is, if the DET were indeed in construction with a larger N consisting of N and S rather than with the smaller N, we would predict not only that (3.58a) reduces to (3.58b), but also that (3.59a) does not reduce to (3.59b). As can be verified, both predictions are incorrect.

(3.59) a. I heard John's claim that women are equal to men but not Mary's claim that men are inferior to women.
b. I heard John's claim that women are equal to men but not Mary's that men are inferior to women.

A fourth undesirable consequence of accepting Chomsky's analysis is that the extraposition of relative clauses and of sentences in apposition could no longer be stated as one rule; instead, we would have to formulate a rule of Extraposition-from-N, in addition to Extraposition-from-NP. This is so because it is necessary to specify the category of the constituent that the extraposable clause is in construction with, or we would predict extraposition in cases like (3.60).

(3.60) a. Although I hate *the idea, that Bill will have to go to prison* is something we will have to face, sooner or later.
b. As I once told *a little boy, who steals my purse* steals trash.

We have thus seen four reasons for believing that constructions containing sentences in apposition are properly analyzable into two constituents from which we may conclude that they satisfy (3.1ii). We now attempt to show that such constructions satisfy (3.1i) by showing that they satisfy (3.1iii).

As far as I am aware, NP's which allow sentences in apposition do not require them. Thus, *a hypothesis*, *the hypothesis*, *that hypothesis* and *Bill's hypothesis* can all occur as independent NP's without modifying clauses. True, (3.61a) is preferable to (3.61b), but the latter seems to me no worse than (3.61c). (3.61b) and (3.61c) are strange without sufficient context, because *the* implies a known referent, but does not make unique identification possible.

(3.61) a. I heard the claim that the earth is round.
b. I heard the claim.
c. I spoke to the boy.

On the other hand, some of the abstract nouns which allow sentences in apposition disallow certain determiners, or disallow sentences in apposition with certain determiners. Thus, we have neither **Bill's fact* nor **Bill's fact that Mary is a thief* (for semantic reasons, presumably since it is hard to see what such constructions can mean). Specific combinations of determiners and modifying sentences are also bad, as seen in (3.62).

(3.62) I am aware of $\begin{matrix}\text{(a)}\\ \text{(b)}\\ \text{(c)}\end{matrix} \left\{\begin{matrix}\text{*a}\\ \text{the}\\ \text{*that}\end{matrix}\right\}$ *fact* that the earth is round.

Both (3.62a) and (3.62c) imply that there is more than one fact that the earth is round, and are therefore out on semantic grounds. Notice that (3.62c) is excluded by a principle we have already considered, namely, that narrower specification of an already necessarily unique referent is contradictory. Notice that if we substitute *claim* for *fact* in (3.52), all the readings become acceptable; the reason is that *claim*, unlike *fact*, is ambiguous between the content of a claim and an instance in which that claim was expressed, and the latter reading is possible with *a* and *that*, for there can be more than one occasion on which a claim is given expression. *Hypothesis* is not ambiguous in the way *claim* is, but it can be safely substituted for *fact* in (3.62), because there can be more than one hypothesis claiming that the earth is round.[7]

---

[7] There are apparent counterexamples to my hypothesis that a unique referent cannot be more narrowly specified with preservation of acceptability. Thus, (i) is acceptable, although the sun is a unique object and nevertheless *sun* carries the determiner *that*.

(i) Every time I look at that cruel sun, I curse it.

The explanation is, however, quite straightforward: preposed adjectives like *great* in *the great man* are ambiguous between a restrictive and a non-restrictive reading, and properties denoted by nouns are ambiguous in the same way. Thus, if *a man* is analyzed at some level as *a X which is a man* (as suggested in Bach (1968)), the relative clause which predicates a property of X could be either restrictive or non-restrictive. I suggest that the acceptability of (i) can be explained by analyzing *that cruel sun* as *that X, which is cruel, and which is a sun*, an analysis which seems to me correct on intuitive semantic grounds as well.

3.2.3.2. having shown that nouns with sentences in apposition satisfy (3.1) we shall now show that they satisfy (3.2).

(3.2i) is illustrated in (3.63).

(3.63) a. I heard the claim that Bill is mad.
b. *What did I hear that Bill is mad?

As was explained in connection with Nuclei of relative clauses, (3.2ii) cannot be illustrated with the structure $[^{NP}_{NP}$ S], since the maiming of the constituent $[^{DET}_{NP}$ N] is independently prevented by (3.2i) and the Left Branch Condition (in English).

(3.2iii) is illustrated in (3.64).

(3.64) a. The claim that Bill is mad is odd.
b. The claim is odd that Bill is mad.
c. A hypothesis that the number of primes is infinite was proposed.
d. A hypothesis was proposed that the number of primes is infinite.

(3.2iv) is illustrated in (3.65).

(3.65) a. I heard the claim that Bill loves Mary.
b. *Who did I hear that claim that Bill loves?

3.2.4. We have considered three kinds of complex NP's in this section so far, and have argued that they satisfy both the requirements in (3.1) and the predictions in (3.2). There is, however, a kind of construction which, albeit not a *complex* NP, is nevertheless a N & S;[8] however, this N & S-type is not subject to (3.2iv). Thus, consider (3.66).

---

[8] That the construction I am about to introduce is a N & S (i.e., that it satisfies (3.1)) can be shown by using the same kind of arguments I have used for NP's modified by relative clauses or complement sentences. The argumentation in this case raises no new problems as compared with the already considered cases, and I leave its construction to the reader.

(3.66) a. I showed you a picture of an owner of that island.

b. That island, which I showed you a picture of an owner of, is quite large.

In deriving (3.66b) from a structure which contains (3.66a), the triply-underscored NP has been chopped. As the chopped element is the Satellite of the doubly-underscored NP, its chopping is, of course, allowed by (3.2iii), but since the doubly-underscored NP is in turn the Satellite of the singly underscored NP, (3.66b) in effect violates (3.2iv), since it arises through the maiming of the Satellite of the singly-underscored NP.

Construction like (3.66a) were studied by Ross in connection with the optionality of Pied-Piping, and he pointed out that no S-node can intervene between the NP to be chopped and some higher NP which dominates it.[9]

Such constructions have the distinguishing property that they can be generated by the recursive application of one rule, such that it embeds an NP as complement to a head NP. In other words, given the chopped element, the structure of the highest N & S to which it belongs is entirely predictable (except for the internal constituency of the various head NP's) by one rule. While this property sheds little light on the peculiar behavior of the structures which possess it and would have to be stated as an ad hoc rider on (3.2iv) so long as the latter is regarded as a formal constraint, I shall argue in Chapter Four that if (3.2iv) is regarded as a performance constraint, cases like (3.66) no longer constitute counterexamples; within that framework, the inner predictability of the structures at issue will emerge as a significant distinguishing property.

3.3. This section is broken down into two main subsections devoted to a discussion of optional adverbials modifying sentences and verb phrases respectively. That at least those two classes of adverbials should be recognized is a fairly well-established point; thus, it has often been pointed out that only S-adverbials can be preceded by pause or preposed, while only (some) VP-adverbials may undergo Passive or belong to antecedents of do so. Additional reasons for recognizing this distinction were given in (1.118g), and I repeat them below in some detail.

---

[9] Ross introduces the further restriction that coordinate NP-nodes cannot be reordered, in order to prevent (ii) from being derived from (i).

(i) I watched Bill and the boy.
(ii) *The boy Bill and whom I watched was vain.

Notice however, that the added condition is unnecessary, since wh-marking is a feature-changing rule prevented from going into coordinations in any case.

First, it is a generally accepted assumption that two adverbs of the same kind cannot occur in the same clause. Therefore, the existence of grammatical sentences like those in (3.67) suggests that either the singly and the doubly underscored strings in each example are not adverbs of the same type or that they do not belong to the same clause; I think that only the latter position is tenable.

(3.67) a. John$_j$ slapped his kid$_i$ because he$_i$ had been too quiet, because his$_j$ wife asked him$_j$ to do that.
b. John went to the store in order to buy some whiskey, in order to antagonize his mother-in-law.

The conclusion that the doubly underscored strings belong to higher sentences than the singly underscored ones seems to me rather unavoidable.

Second, one condition on neutralization to which I know of no exceptions is that the neutralizer must command the neutralizee (if not in surface structure, at least at some (usually later) point in the derivation). Most informants agree that (3.68a) is considerably more acceptable than (3.68b).

(3.68) a. John didn't beat up his wife because he ever enjoyed any of that.
b. *John didn't beat up his wife, because he ever enjoyed any of that.

The distinction in grammaticality between (3.68a) and (3.68b) can be readily explained if we assume that in the former, but not in the latter, the neutralizer commands the neutralizee, and failure of command can be achieved in (3.68b) only if the adverbial belongs to a higher clause than the main one.

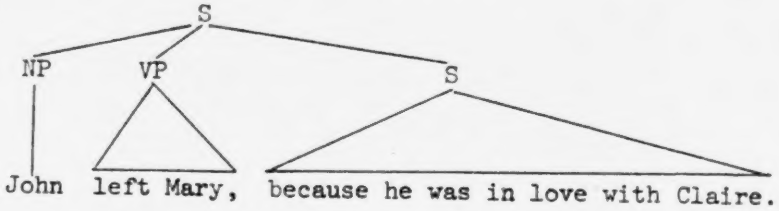
Third, the neutralization of comparees by comparators can only be accounted for by assuming that comparative and equitative clauses are VP-modifiers. The grammaticality of (1.66a) and (1.66b) and the ungrammaticality of (1.66c) and (1.66d) shows that such neutralization must be recognized.

3.3.1.1. Concerning the behavior of S-adverbials with respect to (3.1), if it can be shown that the underlined string in (3.69) is a constituent of type S, it will thereby have been shown that (3.69) satisfies (3.1i) and (3.1ii).

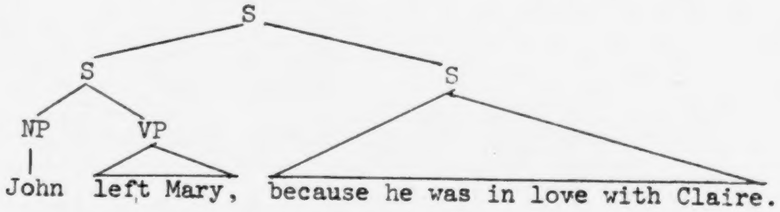
(3.69) John left Mary, because he was in love with Claire.

Notice that the first two arguments given in section 3.3 hold only if we assume that (3.69) has the structure schematically represented in (3.71), but not that represented in (3.70).

(3.70)



(3.71)



Indeed, only given a structure like (3.71) do the two adverbs in (3.67a) or (3.67b) belong to separate clauses, and only then does Negation fail to command *ever* or *any* in (3.68b). But if (3.71) rather than (3.70) is the correct representation of (3.69) then the underlined string in the latter is indeed a constituent (that it is a constituent of type S hardly requires argumentation, since what else could it be?).

As I pointed out in (1.118g), Ross offered two arguments of his own in defense of the position that (3.69) involves S-, rather than VP-modification. I review them below.

Ross first argument was that the pause indicated by comma in (3.69) is not obligatory, but if we pause at all, the natural place for pausing is after the main clause, rather than after *John*; this, he claimed, shows that the immediate constituents of (3.69) are the main clause and the adverbial, rather than *John* and the remainder of (3.69). This argument is unconvincing because: (i) pauses between (short subjects and VP's are quite bad in general, as shown by (3.72),

(3.72) *John, left Mary.

and (ii) if we do not pause after the main clause in (3.69), it is no longer *necessary* that the adverbial be a S-modifier, as the singly underscored string in (3.67b) shows.

Ross' second argument is directed against Langacker (1969), where it is claimed that the ungrammaticality of (3.73c) is due to the requirement that extraposed clauses be commanded by the constituent they cross. Ross proposes that the paradigm in (3.73) be accounted for by his RRC, since Extraposition can also operate from object, in which case no constituent is crossed.

(3.73) a. That John is free is surprising because criminals belong in jail.
b. It is surprising that John is free because criminals belong in jail.
c. *It is surprising because criminals belong in jail that John is free.

Ross concludes that the adverbial in (3.73a) must have the structure represented in (3.71) for the RRC to yield the correct results.

As I pointed out in (1.118g), this argument does not really go through, for extraposed clauses can cross non-clausal adverbials, as shown in (1.126c). The unacceptability of (3.73) is probably due to the fact that rightward movements are fairly bad when crossing a sentence, whether the latter is an adverbial or not; thus, (3.74) is no better than (3.73), although no adverbial is involved.

(3.74) *I looked the desk which I had received from my brother-in-law over.

An additional reason for rejecting this argument of Ross' is provided by (3.75).

(3.75) a. A girl who was sick just left because a storm was raging outside.
b. *A girl just left because a storm was raging outwise who was sick.

(3.75), unlike (3.73a), is ambiguous between a sentential and a verb phrase reading of the adverbial. If Extraposition-from-NP were constrained by the RRC, as Ross claims, we would expect (3.75b) to be ill-formed on the S-modification reading, but well-formed on the VP-modification reading; the fact that (3.75) has no well-formed readings falsifies Ross' hypothesis.

I conclude that the above two arguments of Ross' do not go through; nevertheless, I have given other arguments for believing that the position he was trying to defend was essentially correct.

Having shown that S-adverbials satisfy (3.1i) and (3.1ii), it remains to show that they satisfy (3.1iii). It seems to me that the optionality of S-adverbials is intuitively as clear as that of non-restrictive relatives. But since I have already pointed out that foolproof operative tests are, in all probability, impossible to find, I shall not press the matter any further. Apparent counterexamples like (3.76) may be handled by requiring that some features of Tense and/or Aspect form a component with specific adverbials in underlying representation (see McCawley (1971)), and that the discontinuous component consisting of Perfect and of the time adverb in (3.76) be regarded as an optional dependent of the main clause.

(3.76) John has been in Columbus since the beginning of the term.

3.3.1.2. That the behavior of S-adverbials confirms the predictions in (3.2i) -(3.2iv) is shown by (3.77)-(3.80), respectively.

(3.77) a. That John left, because he was sick, is obvious.
b. *What because he was sick is obvious is that John left.
c. *(It) because he was sick is obvious is that John left.

(3.78) a. John left the city, because he was sick.
b. The city which John left, becauee he was sick, is a real megalopolis.

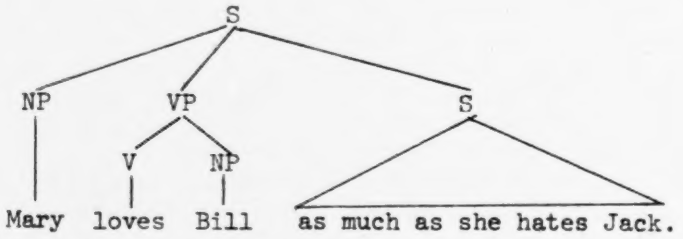
(3.79) a. Mary thinks John left the city, because he couldn't stand the atmosphere.
b. It's because he couldn't stand the atmosphere that Mary thinks John left the city.

(3.80) a. John found Mary at the place of the accident.
b. *The accident which John found Mary at the place of was a tragic one.

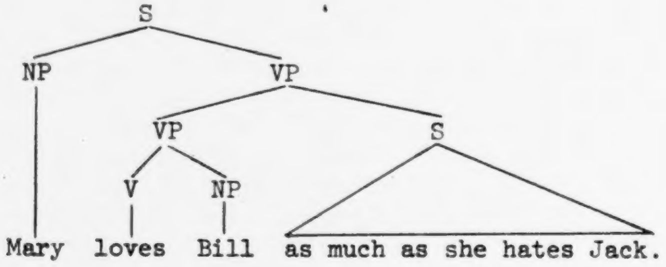
3.3.2.1. In showing that optional VP-adverbials satisfy (3.1), the main problem is to show that the structure of (3.81) is essentially (3.83) rather than (3.82).

(3.81) Mary loves Bill as much as she hates Jack.

(3.82)



(3.83)



To begin with, it seems to me that the facts of optionality are no more obscure for certain VP adverbs than they are for S-adverbs in general. Thus, it seems intuitively clear that the adverbial clause in (3.81) is optional, while the adverb in (3.84) is not, the distinction is less obvious, however, for (3.85), where the instrumental adverb is, I claim, optional on the surface, but not in underlying structure.

(3.84) John went to school.

(3.85) John cut some bread (with a knife).

There is, however, one fact which suggests that the structures of (3.81) and (3.85) are different: optional adverbs are subject to (3.2iv), but obligatory ones are not, as the contrast in acceptability between (3.86) and (3.87) shows.

(3.86) a. John cut the bread with an edge of his sword.
b. ?What did John cut the bread with an edge of?

(3.87) a. John left the city in spite of the owner of the palace.
b. *What did John leave the city in spite of the owner of?

This fact can be made sense of if the structure of (3.81) is (3.83), but not if it is (3.82). This, together with intuitions about optionality, support my claim that optional VP-adverbials satisfy (3.1).

3.3.2.2. That optional VP-adverbials satisfy (3.2i)-(3.2iv) is shown by (3.88)-(3.91) respectively.

(3.88) a. That John loves Mary as much as Jack hates Jill is obvious.
b. *What as much as Jack hates Jill is obvious is that John loves Mary.
c. *(It) as much as Jack hates Jill is obvious that John loves Mary.

(3.89) a. John loves Mary more than Jack hates Jill.
b. It's Mary who John loves more than Jack hates Jill.

(3.90) a. I suspect John loves Mary as much as Jack hates Jill.
b. It's as much as Jack hates Jill that I suspect John loves Mary.
c. John likes Bill more than the owner of the house.
d. It's the owner of the house who John likes Bill more than.[10]

---

[10] An intriguing fact, for which I have no explanation, is that (iia) and (iib) are unambiguous, although the source, (ia), is ambiguous between (ib) and (ic).

(i) a. John likes Mary more than Bill.
b. John likes Mary more than Bill likes Mary.
c. John likes Mary more than John likes Bill.

(ii) a. ?It's Bill who John likes Mary more than.
b. ?Who does John like Mary more than?

Indeed, (iia) or (iib) can only be derived from (ic), where the constituent to be moved is an object, but not from (ib), where it is a subject. This is extremely puzzling, for there is no general principle preventing the movement of subjects, as (iiib), derived from (iiia) shows.

(iii) a. John is taller than Bill (is tall).
b. It's Bill who John is taller than.

(3.91) a. John likes Mary as much as Bill hates Jill.
b. *It's Jill who John likes Mary as much as Bill hates.
c. John likes Bill more than the owner of the house.
d. *It's the house that John likes Bill more than the owner of.

3.4. This section is devoted to coordinate structures, and is broken down into two main subsections devoted to asymmetric and symmetric coordination respectively.

Various writers (Lakoff and Peters (1968), Zwicky (1969), McCawley (1971), R. Lakoff (1971)) have pointed out and discussed the distinction between symmetric coordination (in which the terms are reversible salvo sensu) and asymmetric coordination (in which the terms are not so reversible). Asymmetric coordination has been defined in a narrow and in a broad sense. The narrow sense (which is the one proposed by Zwicky, if I understand his account correctly) concerns instances of coordination which are necessarily asymmetric, such as (3.92a), which cannot be understood synonymously with (3.92b).

(3.92) a. John was shot in the chest and then died of that wound.
b. John then died of that wound and was shot in the chest.

A subset of necessarily asymmetric coordinations consists of those which cannot be reversed with preservation of grammaticality as shown in (3.93).

(3.93) a. John went and solved the problem.
b. *John solved the problem and went.

The broad sense of 'asymmetric coordination' (proposed by R. Lakoff) concerns any coordination in which an interpretation of causality or temporal order between the terms can be imposed, even by using extra-linguistic information.[11] Thus (3.94) is regarded by R. Lakoff as either an instance of symmetric conjunction of poor acceptability, or an instance of asymmetric conjunction based on the assumption that John owns a Ford.

---

[11] R. Lakoff says that, in asymmetric coordination, the first term is not only asserted, but also presupposed by the second. However, this use of "presuppose" is rather misleading, for "presupposition" has been generally used in the literature to refer

(3.94) Fords can go fast, and John just got a ticket for speeding.

R. Lakoff is careful to point out that the predicate CAUSE or PRECEDE holds between (sometimes very elaborate) inferences which can be made from the coordinate terms, and not simply between the coordinate terms, as suggested in Postal (1971). That the latter proposal cannot be correct is shown by asymmetric coordinations like (3.95).

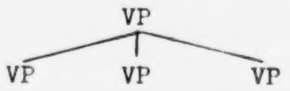
(3.95) The boy who cursed him is dead now, and John was revoltingly happy throughout the morning.

Clearly, it is absurd to assume that the fact that the boy in question is dead now preceded John's feelings of happiness in the morning; rather, the event which preceded and caused John's happiness is "the boy who cursed John died", and the latter can only be inferred from the first conjunct in (3.95).

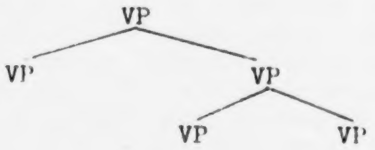
The various writers who have concerned themselves with coordination noticed that it is not easy to distinguish between symmetric and asymmetric coordination formally. The only difference in surface structure (pointed out by R. Lakoff) is that, given a sequence of more than two coordinate terms, the latter must be grouped by two if they are asymmetric, while this is not necessary when they are symmetric. In other words, given a string like (3.96), its phrase structure can be either (3.97) or (3.98) on a symmetric interpretation, while it must be (3.98) on an asymmetric one.

(3.96) Every day, John eats, drinks and goes to work.

(3.97) .....



(3.98) .....

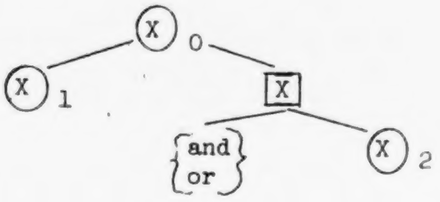


However, this distinction is neutralized in binary coordinations, whose surface structure is (3.99), regardless of symmetricity.

---

to those assumptions definable on purely linguistic grounds; a more adequate term in this case might be "invited inference", as defined in Zwicky and Geis (1971). For a discussion of some sloppy uses of the term "presupposition", see Garner (1971).

(3.99)



The existence of the boxed-X-constituent was convincingly defended in Ross (1967), but I think there are good reasons for regarding the boxed X as different in kind (in the sense of Appendix Two) from the circled X's, the main reason being that the boxed one cannot occur alone--whether it is a S, a NP or a VP--while the circled ones can.

The interesting feature of (3.99) for our purposes is that both kinds of coordination appear to be N & S's. Indeed, X_0 directly dominates X_1, both nodes belong to the same category, and the boxed X is optional, because X_0 and X_1 are categorical nodes of the same kind. Moreover, since the boxed X cannot occur alone, there is a principled reason for regarding X_1 as the Nucleus and the boxed X as the Satellite of X_0.

I am well aware that my claim that coordinate structures are N & S's is rather startling, as it flies in the face of most syntactic analyses, which consider coordination and subordination as entirely distinct and unrelated structures. In the remainder of this chapter, I shall try to show that my proposal is more than mere terminological jugglery, made possible, perhaps, by insufficiently strong constraints placed on the definition of N & S by (3.1). It should be clear, however, that I do not claim that coordination and subordination are indistinguishable, but merely that they share more properties than meet the eye. Thus, my claim that coordinations are N & S's is not different in spirit from the claims made by a number of writers to the effect that verbs, adjectives, prepositions, etc., all belong to the larger category 'predicate'; obviously, none of these writers meant that no distinction is necessary between the various kinds of predicates, but only that their common features must be recognized.

3.4.1.1. We begin our discussion with asymmetric coordination. I suggested earlier that in (3.99), X_1 should be regarded as the Nucleus and the boxed X as the Satellite because the former can occur alone while the latter cannot. However, we need to find stronger criteria for identifying the Nucleus and the Satellite, since (a) there are languages without an overt morpheme and, in which conjunction is expressed by juxtaposition; (b) the first coordinate term may be introduced by both or either, and it is no longer obvious then that it is optional; (c) (3.1iii) requires that optionality be proven with respect to both surface and underlying structure, not just the former. In view of these facts, it seems preferable to choose the Nucleus and the Satellite between X_1 and X_2, rather than between the former and the boxed X (in keeping with the suggestion made in Chapter One that adverbials--and, by extension, dependent terms in general--be ambiguously defined in such a way as to include or exclude the connector).

Thus, if we confine our choice to X_1 and X_2, we find that there are semantic reasons for regarding the former as the Nucleus and the latter as the Satellite in asymmetric coordination, since the latter, but not the former, is semantically optional. This is so because X_1 or its negation is 'presupposed' (in the sense of R. Lakoff [see footnote 1]) by X_2 in asymmetric conjunction or disjunction respectively, while X_2 is not presupposed by X_1. To see this, consider (3.100) and (3.101).

(3.100) Two and two are four, and I'll fail you for writing that two and two are seven.

(3.101) You will leave my room, or I'll shoot you.

The meaning of the second conjunct in (3.100) is incomplete without the presupposition that two and two are four, for, without it, my decision to fail you appears arbitrary (we must assume an arithmetically naive speaker-listener, who needs to hear the first conjunct in order to be aware of the presupposition). Similarly, if I utter only the second disjunct in (3.101), part of its meaning will be lost, for it is not the same to make the threat that I will shoot you and to say that I will do so only if you don't leave my room (the biconditional inference is in fact an invited one, as shown in Geis and Zwicky). In contradistinction, the meaning of the first term is complete without the second, for two and two are still four, even if you write that they are seven on a test, and the only thing I am asking you to do in (3.101) is to leave the room, regardless of what I intend to do in case you don't. The point can perhaps be made clear by considering the analogy of antecedents and pro-forms. A pro-form presupposes an antecedent, but an antecedent does not presuppose a pro-form. Thus, in *John told Mary that he had a son*, part of the meaning of *he* is lost if we don't know it stands for *John*, while no part of the meaning of *John* is lost if we don't know that the same person is later referred to with *he*. Thus, the optionality of the second term in asymmetric coordination (X_2 or the boxed X in (3.99)) suggests that this type of construction is similar to a main clause modified by an adverbial clause.

A second reason for regarding second asymmetric coordinate terms as similar to adverbials is provided by the behavior of backwards pronominalization. Thus G. Lakoff (1968) noticed that backwards pronominalization can proceed from adverbial clauses into main ones if the pronoun is a non-subject, but not if it is a subject, as (3.102) and (3.103) show. However, as the distinction in acceptability between (3.103a) and (3.103b) shows, the length of the second term is important, as it determines the level of phonetic stress on the antecedent, and the result is acceptable only if the antecedent is weakly stressed. On the other hand, as can be seen in (3.102a) no amount of stress reduction can save backwards pronominalization into a subject.

(3.102) a. *He$_i$ was hit by Mary, before J$\breve{o}$hn$_i$ had the slightest chance to get up and leave.
b. *He$_i$ was hit by Mary, before J$\hat{o}$hn$_i$ left.

(3.103) a. Mary hit him$_i$, before Jŏhn$_i$ had the slightest chance to get up and leave.
b. *Mary hit him$_i$, before Jôhn$_i$ left.

The interesting fact is that asymmetric conjunction behaves exactly like main clauses modified by adverbial clauses with respect to backwards pronominalization, as the perfect parallelism between the paradigm in (3.104), (3.105) and that in (3.102), (3.103) shows (it is, of course, necessary to put an asymmetric interpretation on the sentences in (3.104) and (3.105).

(3.104) a. *He$_i$ was killed by Mary, and Jŏhn$_i$ was buried with full honours in the family mausoleum.
b. *He$_i$ was killed by Mary, and Jôhn$_i$ was buried.

(3.105) a. Mary killed him$_i$, and Jŏhn$_i$ was buried with full honours in the family mausoleum.
b. *Mary killed him$_i$, and Jôhn$_i$ was buried.

The unacceptability of (3.104a) can only be explained if we assume that a first asymmetric conjunct is in some sense similar to a main, rather than a subordinate, clause. It follows therefore that the suggestion made in Postal (1971) to the effect that asymmetric coordination is derived from subordination, in the sense that the terms are arguments of some predicate like PRECEDE or CAUSE, is definitely wrong. This is shown quite clearly by (3.106), since, if the b-sentence were indeed derived from the a-one, there should be no difference in acceptability between the two.

(3.106) a. That she$_i$ was shot in the chest brought it about that Mary$_i$ eventually died.
b. *She$_i$ was shot in the chest, and Mary$_i$ eventually died.

A third reason for regarding asymmetric coordinate structures as N & S's is their behavior under maiming. Indeed, maiming a first asymmetric coordinate term is infinitely more acceptable than maiming a second one, as predicted by (3.2ii) and (3.2iv), and as shown by (3.107)-(3.109).[12] The contrast between the b- and the c-sentences in (3.107)-(3.109) can be explained neither by the CSC, nor by Ross' claim that asymmetric coordination is exempt from the CSC.

---

[12] If some of the b-sentences are not fully acceptable to all speakers, this is probably due to the fact that there are more explicit and less ambiguous ways of indicating subordination than through asymmetric coordination (where subordination is usually discovered through elaborate inferences from extralinguistic assumptions), and the sentences in (3.107)-(3.109) violate to some extent Grice's conversational principle that statements should be as clear as possible; in other words, it may be harder for a hearer to realize that the sentences in (3.107)-(3.109) involve N & S's then in cases containing explicit marks of subordination.

(3.107) a. John murdered the girl, and her father (subsequently) suffered a heart attack.
b. ?The girl who John murdered, and her father (subsequently) suffered a heart attack, was my cousin Alice.
c. *The heart attack, which John murdered the girl, and her father (subsequently) suffered, was a fatal one.

(3.108) a. Mary is beautiful, and Bill is happy to be married to her.
b. ?Beautiful though Mary is, and Bill is happy to be married to her, I definitely like Sheila better.
c. *Happy though Mary is beautiful, and Bill is to be married to her, I definitely like Sheila better.

(3.109) a. Bill will have to eat the sandwich, or his mother will tell uncle Joe.
b. The sandwich which Bill will have to eat, or his mother will tell uncle Joe, is a fairly large one.
c. *Uncle Joe, who John will have to eat the sandwich, or his mother will tell, is a rather unpleasant fellow.

A fourth reason for regarding asymmetric coordination as N & S's is that they obey the following principle:

(3.110) If an element of a Satellite is removed from the containing N & S, that element must also exist in and be removed from the Nucleus (under certain conditions which remain to be made precise).[13, 14]

---

[13] An explanation for (3.110) and a refutation of an implied counterproposal made in Neubauer (1970) are offered in Appendix Four.

[14] I have not investigated in detail the exact conditions under which across-the-board maiming can occur. However, it seems clear that the identical elements must be in 'similar positions', since two elements cannot be subject and object respectively, as shown by (i).

(i) a. John$_i$ likes Mary and (therefore) Bill hates him$_i$.
b. *It's John who likes Mary and (therefore) Bill hates.

On the other hand, the two elements need not be objects of the same kind, as demonstrated by (ii).

(ii) a. John loves Mary$_i$ but can never agree with her$_i$.
b. It's Mary who John loves but can never agree with.

If across-the-board maiming depends on the parallelism of the positions of the moved elements, the parallelism of the larger structures from which the elements in question are removed is also a pertinent factor. Thus, while across-the-board maiming of coordinate structures of any kind seems to be universally acceptable (subject to the restriction mentioned at the beginning of this note), across-the-board maiming of matrices and complements, as in (iii), are marginal to some speakers, and so are (3.111d) and (3.112d).

(iii) a. John told Mary that he loves her.
b. It's Mary who John told that he loves.

The marginality of these sentences is presumably due to the fact that the structures containing the identical elements do not belong to categories of exactly the same kind. When the containing structures belong to clearly distinct categories, as in (iv), where the categories are S and NP respectively, the result is considerably worse.

(iv) a. The doctor cured John in spite of himself.
b. ?*It's John the doctor cured in spite of.

A strange effect of parallelism is that it seems to improve even structures which would otherwise be unacceptable. (v) shows that violations of the CNPC are less severe when they occur across-the-board.

(v) a. *It's John who I believe the claim that you killed.
b. ?*It's John who I believe the claim that you killed and disbelieve the allegation that you tried to help.

(3.110) is a formalization of an observation made in Ross (1967, sec. 4.2), and illustrations are provided in (3.111) and (3.112).

(3.111) a. John likes Mary$_i$ more than Bill hates $\left\{\begin{matrix}\text{her}_i \\ \text{Jill}\end{matrix}\right\}$.

(3.111) b. It's Mary$_i$ John likes more than Bill hates $\left\{\begin{array}{l}\text{her}_i\\ \text{Jill}\end{array}\right\}$.
c. *It's $\left\{\begin{array}{l}\text{Jill}\\ \text{her}_i\end{array}\right\}$ John likes Mary$_i$ more than Bill hates.
d. It's Mary John likes more than Bill hates.

(3.112) a. John tore the rug$_i$ in rolling $\left\{\begin{array}{l}\text{it}_i\text{ up}\\ \text{up the curtain}\end{array}\right\}$.
b. It's the rug, John tore in rolling $\left\{\begin{array}{l}\text{it}_i\text{ up}\\ \text{up the curtain}\end{array}\right\}$.
c. *It's $\left\{\begin{array}{l}\text{the curtain}\\ \text{the rug}_i\end{array}\right\}$ John tore $\left\{\begin{array}{l}\text{the rug}\\ \text{it}_i\end{array}\right\}$ in rolling up.
d. It's the rug John tore in rolling up.

That asymmetric coordination behaves in the same way as sentences modified by adverbial clauses with respect to (3.110) can be seen by comparing (3.113), (3.114) with (3.111) and (3.112).

(3.113) a. John murdered Alice, and Bill swore to avenge her.
b. Alice, who John murdered, and Bill swore to avenge her death, was a very nice girl.
c. *It's $\left\{\begin{array}{l}\text{Alice}\\ \text{her}\end{array}\right\}$ John will have to marry $\left\{\begin{array}{l}\text{her}\\ \text{Alice}\end{array}\right\}$ or let go.
d. It's Alice John will have to marry or let go.

3.4.1.1.2. Regarding asymmetric coordination and (3.2), (3.115) shows that (3.2i) is entirely confirmed.

(3.115) a. I told you that John murdered Alice, and her father subsequently suffered a heart attack.
b. *What I told you, and her father suffered a heart attack, is that John murdered Alice.
c. I told you to get out, or I'll hit you.
d. *What I told you, or I'll hit you, is to get out.

The confirmation of the predictions in (3.2ii) and (3.2iv) was discussed in 3.4.1.1.1 and illustrated with (3.107)-(3.109).

The prediction in (3.2iii) is not confirmed in English as there are no convincing examples of second asymmetric coordinate terms being moved, with or without the connector. I do not, however, regard this fact as particularly damaging to my analysis, since instances of 'split conjuncts' have been uncovered in other languages, and, as I pointed out already, there is no guarantee that the option offered by statements like (3.2ii) and (3.2iii) will be taken in each and every language. The instances of split conjuncts which I have come across in the literature involve symmetric conjunction, and will be taken up in section 3.4.2.2. I see no reason for believing that conjunct-splitting is limited to symmetric conjunction, but since I do not know of any discussion involving asymmetric conjunct-splitting in other languages (particularly when asymmetry is not decidable by strict linguistic criteria) I will not pursue the matter any further.

The absence of split conjuncts (and coordinate terms in general) in English is, I believe, due to a phenomenon of contamination of the restrictions on symmetric *and* and *or* to all the forms which superficially resemble those coordinating conjunctions. Thus, there are instances of *and* which, on strictly linguistic grounds, are obviously not conjunctions (see also 3.4.1.2), such as in (3.116a), which is in all probability related to, or even derived from, (3.116b); nevertheless, the constituent introduced by *and* cannot move, as shown in (3.117a), unlike the constituent introduced by *to*, which can, as shown in (3.117b).

(3.116) a. You will have to try *and* finish your paper.
b. You will have to try *to* finish your paper.

(3.117) a. *What you will have to try is and finish your paper.
b. What you will have to try is to finish your paper.

The behavior of this kind of *and* parellels symmetric *and* with respect to copying as well; thus (3.118a) is marginal compared with (3.118b).

(3.118) a.??What you will have to try and do is finish your paper.
b. What you will have to try to do is finish your paper.

(3.118a) seems to have the same degree of acceptability as (3.119b), where we are dealing with symmetric *and*; the restriction is probably the same as that which marks as marginal conjunctions of indefinite and definite forms in general, as shown in (3.120).

(3.119) a. Mary tried to laugh and to dance.
b.??What Mary tried to laugh and to do was to dance.

(3.120) a.??Someone and Bill walked in.
b.??A cat and the dog ran away.

It is interesting that this contamination phenomenon does not extend to *but* in the sense of *except*; thus, (3.121) strikes me as fairly tolerable.

(3.121) a. You've been talking about nothing but John.
b. John, you've been talking about nothing but.

In conclusion, even though the absence of confirmatory evidence for (3.2iii) is an embarrassing fact, it is probably not a genuine counterexample.

3.4.1.2. Before leaving the topic of asymmetric coordination, we shall consider a subclass of asymmetric conjunction which apparently violates (3.2iv), and which led Ross to the claim that asymmetric conjunction is exempt from the CSC. The phenomenon is illustrated in (3.122) and (3.123).

(3.122) a. I'll have to try and finish that book before dark.
b. What will I have to try and finish before dark?

(3.123) a. John went to the store and bought some whiskey.
b. It's some whiskey that John went to the store and bought.

I believe, however, that the conjunctions in (3.122) and (3.123) are *not* N & S's, since neither term is semantically optional; thus, we seem to be dealing with two stages of *one event*, so that each term presupposes (in the sense of R. Lakoff) the other. Indeed, in (3.107a), there is no necessary connection between the girl's death and the heart attack her father suffered (John might have committed the murder for the purpose of bringing about the attack, but the sentence is non-specific on that score). In contrast, in (3.123), there is a strong suggestion that the two conjuncts express stages of *the same event*, and both, consequently, presuppose each other; the coordination is thus *not* an N & S, for (3.1iii) is not satisfied.

The suggestion (or invited inference) that we are dealing with one event is due, I think, to the fact that the two conjuncts have the same subject *and* to the fact that the second occurrence of that subject is reduced. Consider the gradual increase in acceptability exhibited in the following paradigm:

(3.124) a. *This is the whiskey which John$_i$ went to the store and Paul bought.
b.?*This is the whiskey which John$_i$ went to the store and hê$_i$ bought.
c. ?This is the whiskey which John$_i$ went to the store and hě$_i$ bought.
d. This is the whiskey which John went to the store and bought.

In (3.124a), the subjects are different; in (3.124b), the subjects are the same, but the second one is unreduced; in (3.124c), the second cubject is partly reduced, while in (3.124d), it is totally reduced. This paradigm shows that the stronger the invited inference to the effect that the two conjuncts refer to parts of the same event, the easier it is to maim the second.

It is noteworthy that mutual presupposition--as in (3.124d)--has the same effect on certain purpose adverbs as on second conjuncts: both cease to be Satellites (and therefore become maimable). Thus, (3.125) shows the same increase in acceptability as (3.124).

(3.125)[15] a. *This is the whiskey which John went to the store in order that Paul may buy.
b.*?This is the whiskey which John went to the store in order that hê may buy.
c.??This is the whiskey which John went to the store in order that hě may buy.
d. This is the whiskey which John went to the store in order to buy.

Just as the ideal situation for conjunct-maiming is one in which the conjuncts are VP's, the ideal situation for purpose-adverbial maiming is one in which the adverb is a VP-modifier. To see this even more clearly, contrast (3.125d) and (3.125e) where the purpose adverbials are VP- and S-modifiers respectively.

(3.125) e. *This is the whiskey which John went to the store, in order to buy.

---

[15] Poutsma (1929) and Ross (1967) have in fact proposed that (i) be derived from (ii).

(i) John went to the store and bought some whiskey.
(ii) John went to the store to buy some whiskey.

Later writers (e.g., Zwicky (1969)) have pointed out, however, that (i) and (ii) are not synonymous, for the latter, but not the former, allows the continuation ...but he didn't buy anything.

Another difference between (i) and (ii) was pointed out in Schmerling (1972): only from the latter, but not from the former, is it possible to infer that the going to the store must have been undertaken with the intention of buying the whiskey, as shown by the contrast between (iii) and (iv).

(iii) This is the whiskey which John went to the store and bought, although he had no intention of buying anything.
(iv) *This is the whiskey which John went to the store to buy, although he had no intention of buying anything.

In Partee (1971), it is argued that Conjunction-Reduction is sometimes associated with changes in meaning and the paradigms in (3.123) and (3.124) support her contention. However, (3.125) suggests that the facts Partee discusses hinge on *VP-modification* (as against S-, NP-, etc., modification), rather than on *Conjunction-Reduction*; this is also suggested by (3.126).

(3.126) a. John wants [to go to the store and to buy some whiskey].
b. *This is the whiskey which John wants [to go to the store and to buy].
c. John wants to [go to the store and buy some whiskey].
d. This is the whiskey which John wants to [go to the store and buy].

The difference in acceptability between (3.126b) and (3.126d) suggests that the bracketed constituent in the former is a coordinate *S-node*, while the bracketed constituent in the latter is a coordinate *VP-node*. This suggestion is supported by the fact that (3.126c), but not (3.126a) allows the continuation *and Mary wants to do so too*, as well as by the fact that the pseudo-clefting of the bracketed string in (3.126a) does not leave behind a *do* pro-form (or an underlying action verb *do*, according to Ross (1971)), while the pseudo-clefting of the bracketed string in (3.126c) does, as shown by (3.126e) and (3.126f) respectively.

(3.126) e. What John wants is to go to the store and buy some whiskey.
f. What John wants to *do* is go to the store and buy some whiskey.

The fact that VP-modification behaves alike with respect to coordination or subordination is an additional reason for believing the claim I made earlier that the distinction between these two construction-types is not nearly as hard and fast as it had generally been assumed.

Finally, it should be pointed out that Ross' contention that sentences like (3.122a) and (3.123a) are exempt from the CSC is only partly true. Indeed, the conjuncts can be maimed, but not chopped, as discussed in 3.4.1.1.2 and illustrated in (3.117a).

3.4.2.1. In 3.4, I pointed out that coordinations of either type do satisfy (3.1) and should therefore be regarded as N & S's, with X_1 as the Nucleus and the boxed X as the Satellite. In 3.4.1.1.1, I pointed out that a deeper insight may be achieved by considering the roles of X_1 and X_2, rather than of the former and of the boxed X. In this section, we consider X_1 and X_2 in symmetric coordinations and ask how, given the N & S hypothesis, the Nucleus and the Satellite can be identified on a semantic basis.

From the definition of symmetric coordination, it follows that the relative surface position of X_1 and X_2 is semantically irrelevant, as they can switch places with no effect on meaning; therefore, the fact that X_2 is dominated by the boxed X in (3.99) is without semantic significance. To determine which node is the Satellite, we must ask which one satisfies (3.1iii). The answer is that neither term is optional in an absolute sense, as shown by the syntactic and semantic deviance of (3.127b), but both are optional in a relative sense, as shown by the acceptatility of both (3.127c) and (3.127d); in other words, if one term is present, the other one is optional.

(3.127) a. John and Bill left.
b. *Left.
c. John left.
d. Bill left.

Given the evidence, the only status that can be ascribed to symmetric coordination is the one stated below:

(3.128) Symmetric structures like (3.99) are bidirectional N & S's, in the sense that X_1 is the Nucleus and X_2 the Satellite, and vice versa.

As we shall see in 3.4.2.2, (3.128) will make it possible to explain the syntactic behavior of symmetric cordinations with respect to chopping rules without recourse to additional assumptions. However, I believe that (3.128) is not merely a trick which makes the grammar work, but a statement which makes sense intuitively. (3.128) may give pause at first, as it may seem self-contradictory to call something a Nucleus and a Satellite at the same time; however, Nucleus and Satellite are not incompatible properties, since what makes a constituent a Nucleus is its ability to represent the N & S alone, without further dependents, and what makes it a Satellite is its ability to disappear without altering the construction-type of the larger constituent to which it belongs; as far as I can see, both properties are satisfied by symmetric coordinate terms. It is important to notice that these properties refer to potential, not actual situations; indeed, there is no coordination token in which a term is both present and absent. But surely (3.128) does not assert that any such coordination token does exist. The possible impression of self-contradictoriness derives, I think, from the possible mistaken feeling that (3.128) does in fact make such an assertion.[16]

---

[16] (3.128) can be extended trivially to n-term coordination. Since only one term must be present, but any of the terms could be that term, a symmetric coordination with n terms in a N & S in n ways, such that in each way, some term is the Nucleus and the remainder are Satellites.

I can see at least four reasons for accepting (3.128):

(a) Constructions are either endocentric or exocentric; (3.99) is not exocentric, for X_0 is necessarily of the same type as either X_1 or X_2; therefore, (3.99) must be endocentric.

(b) If (3.99) is endocentric, and we reject (3.128), we may regard both X_1 and X_2 as Nuclei; in other words, symmetric coordinate structures (with n terms) would be 'poly-endocentric'. This hypothesis has some intuition appeal, and in fact represents the traditional position on coordination. However, this position is not without difficulties. First, Nuclei (and Heads in general) are never optional; therefore, endocentric constructions would require an ad hoc statement to the effect that if there is more than one head present, only one is obligatory. Second, the behavior of symmetric coordination with respect to chopping, etc., rules, does not follow from (3.2). Indeed, if symmetric coordinations are pluri-nucleic, no term can move due to (3.2i). But (3.2ii) predicts that Nuclei *can* be maimed, and in order to make the theory conform to the facts, it would be necessary to regard symmetric coordinations as mono-nucleic, the totality of the terms being *the* Nucleus; in that case, 'maiming the Nucleus' could be interpreted as 'maiming all the terms across-the-board, but in no other fashion', which is precisely what we find. This 'solution' is, however, untenable, because the characterization of a construction as simultaneously pluri-nuclei and mono-nucleic is clearly contradictory. Therefore, it would be necessary to include the CSC in the grammar in addition to (3.1) and (3.2). This decision would be unfortunate in the extreme, not only because it would add additional apparatus, but primarily *because it would claim that Nuclei behave differently in N & S's and in poly-endocentric constructions*.

I conclude that (3.128) is to be preferred to the poly-endocentric hypothesis for the reasons expounded above. In addition, it should be noticed that my proposal does *not* really conflict with the traditional view; in fact, it *includes the latter*, since it regards each term as a Nucleus, but it adds the information (which the traditional position misses) that each term is relatively optional, and thus makes possible a principled explanation of the behavior of symmetric coordination with respect to both chopping and maiming.

(c) If we reject both (3.128) and the poly-endocentric hypothesis, it might be proposed that both X_1 and X_2 are dependents of some third element; as no such element exists, this hypothesis is untenable. To propose that both terms are Satellites would make even less sense, for it would then be predicted that both are simultaneously optional.

(d) Finally, it might be proposed that X_1 is the Nucleus and X_2 is the Satellite but not vice versa (as in asymmetric coordination); this would be not only wrong on semantic grounds, but would also make false predictions on the behavior of symmetric coordination with respect to chopping, feature-changing, etc., rules. On the other hand, given (3.128), (3.2) and (3.110) make exactly the correct predictions, as we shall see in the next section.

3.4.2.2. From (3.128) and from (3.2i) and (3.2iv), it follows that neither X_1 nor X_2 in (3.99) can be either chopped or maimed. Thus, the stronger constraints on symmetric coordination, as compared with other construction types, are correctly predicted. The conjunction of (3.128) and (3.2i)-(3.2iv) is almost equivalent to the CSC, but not quite, since the latter predicts that the boxed X in (3.99) is also frozen, while the former makes no such prediction. In fact, my proposal leaves open the possibility that the boxed X may migrate from under X_0.

It is not clear to what extent this possibility is realized in English. Ross (1967) mentions two rules which he regards as counterexamples to his CSC, Conjunct-Insertion and Conjunct-Movement, illustrated in (3.129) and (3.130) respectively.

(3.129) a. John will get a bonus on Friday, and he has been waiting for that since December.
b. John, and he has been waiting for that since December, will get a bonus on Friday.

(3.130) a. John and Bill are similar.
b. John is similar to Bill.

(3.129) is a counterexample only to the strongest form of the CSC, which precludes all movements of nodes dominated by a coordinate node, but not to a weaker form which allows rearrangements within the domination of the coordinate node; it appears that Ross upheld the strong version. (3.129b) does certainly not violate my (3.2), but it does not support my (3.2iii) either, as the latter allows movements of the boxed X even outside the domination of the coordinate node.

(3.130b) does support (3.2iii) as against the CSC; unfortunately, the existence of a transformational relation between (3.130a) and (3.130b) has been cast into serious doubt by Quang (1970), who pointed out that the alleged rule of Conjunct-Movement sometimes fails to preserve meaning and/or grammatical well-formedness or ill-formedness. But even if (3.130a) and (3.130b) are not related transformationally, there are sentences in other languages which suggest that the constituent formed by a second coordinate term plus the coordinating conjunction can be moved.

One such language is Hungarian, where the proposition the teacher sees the boy and a girl can be rendered in any of the three following ways:

(3.131) a tanaar laatja a fiuu eeš ed' laañt
the teacher sees the boy and a girl

(3.132) a tanaar a fiuu eeš ed' laañt laat.
the teacher the boy and a girl sees

(3.133) a tanaar a fiuu laatja eeš ed' laañt
the teacher the boy sees and the girl

I do not know enough about Hungarian to see whether there is enough independent motivation for a Conjunct-Extraposition rule, but notice that the above data are consistent with such a rule, which would derive (3.133) from (3.132), along the lines permitted by (3.2iii).[17]

---

[17] Koutsoudas' account is slightly different: he proposes that (3.133) is derived from (3.131) by fronting the first conjunct. His reasons are that the verb agrees in definiteness with the closest conjunct in both (3.131) and (3.132), and that the fact that it agrees with the first conjunct in (3.133) would require an ad hoc proviso if Agreement applied after the splitting of the conjuncts, when both are equidistant from the verb. However, it seems plausible to expect the possible governors in object-verb agreement to be *a fiuu* and *ed' laant*, rather than the former and *eeš ed' laant*. If so, we may allow the second conjunct to move along with the coordinating conjunction, and to permit Agreement to operate after this movement, for the first conjunct is still closer to the verb than the second conjunct is.

Another language in which the boxed X of (3.99) appears to be movable is Samoan. Grinder (1969) points out that the proposition *the girl cuts the boy with the axe and the knife* can be rendered in one of the following two ways:

(3.134) e sogi e le teine le tama i le to'i *ma le naifi*.
cuts the girl the boy with the axe and the knife

(3.135) e sogi *ma le naifi* e le teine le tama i le to'i.
cuts and the knife the girl the boy with the axe

It seems that the underscored constituent in (3.134) has been moved in (3.135) for topicalization purposes.

I do not know of any other cases in the literature which have been presented as violations of the CSC, but the cases which I know of all involve the movement of the boxed X in (3.99), which supports my (3.2iii), since that node can only be viewed as a Satellite.

If the conjunction of (3.2i), (3.2iv) and (3.128) explains the impossibility of either chopping or maiming one of the terms of a symmetric coordination, the conjunction of (3.2i), (3.2iv), (3.128), and (3.110) explains why across-the-board maiming of symmetric coordinate terms is possible, and why it is in fact the only way in which symmetric coordinations may be maimed. Indeed, since each term is a Satellite of the other, neither can be maimed without the other's being maimed too. The same principle applies to n-term coordinations where any term is a Satellite of every other term (see footnote 16).

I have argued in this chapter that the conjunction of (3.1) and (3.2) improves both observationally and descriptively on the IC, as well as on the island constraints stated in Ross (1967). In the next chapter, I shall argue that my proposal, but not the IC, can receive a natural explanation in terms of general perceptual principles.

CHAPTER IV

INTERRUPTION, CONFLICT, CLOSURE, AND ISLAND CONSTRAINTS

4.0. This chapter is an attempt to provide a plausible explanation of Ross' most important island constraints in terms of principles discussed in Chapter Two and of their partial reanalysis defended in Chapter Three.

The constraint (3.2i) which freezes Nuclei (henceforth: the Frozen Nucleus Constraint) is discussed in section 4.5, where it is shown that violations of it can yield two types of situations which, it is suggested, can be accounted for by a closure and conflict principle respectively.

The constraint (3.2iv) which precludes movements of elements of Satellites outside the scope of the N & S node (henceforth: the Subservient Satellite Constraint) is discussed in section 4.3, where it is argued that violations of this restriction give rise to perceptual conflict.

The Dichotomous Behavior Principle (stated in (1.52)) is discussed in section 4.4, where it is suggested that the distinction between the syntactic phenomena which obey island constraints and those which do not can be accounted for by assuming that the remote relations crucially involved in the two situation-types are recognized by essentially different types of strategies.

The Sentential Subject Constraint is discussed in section 4.2 within the more general framework of the mitigating effect of predictability on the complexity of interrupted behavior. In contrast to Ross' conclusion that the SSC is not universal, it is suggested here that the pertinent constraint (resulting from the lower predictability of the location of a partner in a discontinuous relationship) can apply non-vacuously only in languages where arguments can occur on both sides of the verb, and where moreover one of these positions confers upon some argument a distinguished status (such as in the SVO languages).

In view of the fact that violations of island constraints are not always obvious from a consideration of the surface structure, it is necessary to assume that certain decoding operations take place earlier than others. Section 4.1 examines some experimental and internal evidence which supports the view that some (but not all) decoding operations apply sequentially, and considers the principles which may govern the ordering of such operations. It is argued that the way in which strategies must be ordered for independent reasons is sufficient to account for those instances of island-constraint violations which are not directly observable in surface structure.

4.1. In Chapter Two, I pointed out that the theory of perceptual strategies assumes that the processes it posits are not direct replicae of *the rules of a transformational grammar*, even though it is very likely those processes utilize *transformational information*. In this section, I shall attempt to determine the principles which govern the ordering of strategies, and demonstrate their usefulness for the exploration of specific constraints.

Discussions of ordering have in general been limited to syntactic or phonological transformations; in this respect, the constraining power of *extrinsic rule ordering* (an explicit statement of the order in which rules must apply) has been contrasted with *intrinsic rule ordering* (an inherent property of pairs of rules, to the effect that if both rules apply in a derivation, there is a unique possible order of application dictated by their formulation) (Chomsky (1965)). Later writers pointed out that extrinsic rule ordering is insufficiently powerful to account for newly discovered data, so that ordering paradoxes result unless the grammar is provided with additional apparatus, such as (arbitrary) markers, which code information available at some stage of a derivation for the benefit of later stages of that derivation (as proposed in Kenstowicz and Kisseberth (1970), Postal (1970a), Koutsoudas, Sanders and Noll (1971), Lehman (1971), Baker and Brame (1972) etc.) or, equivalently, global rules, which enable grammatical processes to 'peek' at earlier and/or later stages of derivations (as proposed in Lakoff (1969), (1970b), (1972), Ross (1969b), Postal (1970b), etc.). The proponents of coding or global devices have often argued that since such devices can do whatever extrinsic rule ordering can do but not vice versa, there is little point in having both kinds of devices in a grammar, and they have consequently proposed that extrinsic ordering be discarded from linguistic theory.

Clearly, both extrinsic rule ordering and coding or global apparatus are still a far cry from providing an *explanation* of the empirically discoverable linguistic data. The reason is that the notion 'possible transformation' is at present undefined in linguistic theory, and while a transformational grammar without global devices is too weak in some ways, it is also too strong in other ways; that is, as has frequently been pointed out, it is possible, given the presently available notation, to write a large number of implausible transformations with the same ease with which empirically justifiable ones can be written. Thus, it is not in the least clear that the way in which rules have been formulated is psychologically correct, and since the distinction between intrinsic and extrinsic ordering depends precisely on specific formulations of rules, that distinction itself is up in the air. Similarly, as there is no presently available definition of the notions 'possible coding device' or 'possible global rule', the enormous power of these devices allows, in principle, the prediction of a still larger class of implausible situations.[1]

---

[1] This undesirable increase in power shows that the elimination of extrinsic ordering in favor of global rules is not necessarily a

Most of the effort expended in studying perceptual strategies so far has been aimed at providing proof of their existence as well as adequate formulations of them. Not too much attention has been paid to the order in which strategies must apply, presumably for at least the following reasons: (a) the order of application of specific strategies is not at present amenable to *direct* testing, and (b) the only kind of ordering regarded as interesting in *grammatical* investigations has been the extrinsic one, and an extrinsic ordering of *strategies* appears quite implausible; the reason is that extrinsic ordering in grammars is aimed at limiting the applicability of specific rules, while extrinsic ordering in perception would seem to be perversely aimed at limiting the efficiency of specific decoding operations. In the remainder of section 4.1, I shall try to show that the ordering of perceptual strategies is not without theoretical interest, as even non-extrinsic ordering constrains the class of possible outputs. Specifically, I will consider three syntactic problems which cannot, as far as I can see, be handled with the presently available formal devices without loss of generality, and will argue that a satisfactory solution can be provided by a consideration of the order in which perceptual strategies apply to the pertinent structures .

4.1.1.1. The first problem I shall consider is Ross' (1967) output condition, usually referred to as the NP-over-S Constraint:

(4.1) Internal sentences exhaustively dominated by the node NP are unacceptable.

(4.1) marks as unacceptable sentences like (4.2a) and (4.2b), but does not affect sentences like (4.2c) and (4.2d).

(4.2) a. *It's that Bill may return which disturbs everybody.
b. *It's for Mary to leave that everyone is looking forward to.
c. It's *Bill's playing the piano at all hours* which got the neighbors furious.
d. It's Bill's loud playing of the piano which made his wife ask for divorce.

True, gerundives like the underscored phrase in (4.2c) behave like

---

blessing. If extrinsic ordering is retained and global rules are used sparingly, the number of possible grammars that can account for some set of data is reduced (provided, of course, that we can state in a principled way which classes of phenomena must be described by rule ordering and which by global rules). Chomsky has often argued (e.g., in Chomsky (1970b)) that given the extreme ease with which children learn their languages, a theory which reduces the number of possible grammars applicable to some corpus has a better chance of being psychologically correct than a theory which merely reduces the number of formal devices available to grammars.

sentences with respect to a number of syntactic phenomena, but Ross disposed of this problem by pointing out that constructions like the internal ones in (4.2a)-(4.2d) can be placed on a 'sentencehood-nominality' scale, and that (4.1) is sensitive only to the two highest degrees of sentencehood, rather than to absolute categories. As there are many syntactic phenomena which recognize such a categorial scale (for details, see Ross (in Press)), this explanation can be accepted.

There are, however, at least two problems with (4.1). The first is that it does not always hold; thus, (4.3) seems to me quite tolerable, particularly if it is spoken with a slight pause after the underscored constituent (which is probably what most people would do), even though the internal phrase has the highest degree of sentencehood.

(4.3) The doctor explained that John was sick to the many friends and relatives who had assembled in the hall.

Second, it is not clear at what level (4.1) must hold. Ross regarded (4.1) as an output (i.e., surface structure) condition, but this is clearly incorrect in view of the acceptability of (4.4a) and (4.4b).

(4.4) a. Is where you live known to the police?
b. Isn't whoever steals money a thief?

If we accept the analysis proposed in various places in the literature that free relatives have lexical heads at some level of representation and that embedded interrogatives also originate with heads roughly like the answer to the question, the grammaticality of (4.4a) and (4.4b) is explained if we require that (4.1) hold at the level which precedes the deletion of the heads of the internal clauses. However, the level in question cannot be a well-defined level of linguistic structure. In particular, it cannot be 'shallow structure', if the latter is defined as the output of the cycle, in view of the unacceptability of (4.5a); topicalization is a post-cyclic rule, and the head of the free relative must be present in the string at the point at which Topicalization applies for the latter to be blocked by the CNPC. If shallow structure is defined as the level which precedes all post-cyclic deletions, then (4.1) cannot hold at that level either, for the head must be present when certain deletion rules subject to the CNPC apply, as shown by the ungrammaticality of (4.5b).

(4.5) a. *Bill, I have heard $\left\{\begin{array}{c}\text{the lie which}\\ \text{what}\end{array}\right\}$ Mary told.
b. *This mountain is too massive for me to be able to conceive of $\left\{\begin{array}{c}\text{the object which}\\ \text{what}\end{array}\right\}$ destroyed.

Notice that so far I have assumed that (4.1) is stated at the latest level possible. There is no reason why it could not in principle be stated in shallow structure, even if it is applicable later on as well; however, notice that Subject-Verb Inversion is a post-cyclic rule, and the underscored phrases in (4.4) would not yet be internal at the end of the cycle. One might want to state (4.1) at the level which precedes post-cyclic deletion, but this creates problems of another sort: Kiparsky and Kiparsky (1971) argued that the complements of factive verbs modify the NP the fact at some level. This claim has been recently called into question by a number of writers (I shall return to this matter at the end of (4.1.3.1), but if we accept it for the time being, we can account for the difference in acceptability between (4.6a) and (4.6b) by assuming that the fact is still present in the derivation of the latter at the stage at which the rule which deletes the coreferent of the rock applies. Consequently, the heads of factive complements must be deleted later than shallow structure, exactly like the heads of free relatives.

(4.6) a. This rock is too heavy for me to believe that you intend to pick up.
b.?*This rock is too heavy for me to have realized that you intend to pick up.

However, if we required that (4.1) hold in shallow structure (defined as the level which precedes post-cyclic deletions), we predict that there is a difference in acceptability between (4.7a) and (4.7b) when in fact there is none.

(4.7) a. *Was that John was sick believed by the captain?
b. *Was that John was sick realized by the captain?

The import of the preceding discussion is that (a) at least with respect to factive complements, (4.1) does not hold at a well-defined level of linguistic structure, and (b) the heads of factive complements must be deleted earlier than those of free relatives--for no independent reason--with (4.1) holding, say, in the input to the rule which deletes the heads of free relatives. This solution, while not impossible, seems to me sufficiently ad hoc to suggest that it cannot receive an explanatorily adequate interpretation.

4.1.1.2. The second problem to be considered was raised in 1.2.6, and concerns the fact that certain discontinuities within discontinuities, which are intuitively felt to be instances of the same phenomenon, cannot be stated in the same way in the presently available grammars. Specifically, it was shown that movements out of clauses displaced by Extraposition-from-NP--as in (4.8d)--and movements out of pseudo-cleft phrases--as in (4.9b)--or cleft ones--as in (4.9d)--can only be blocked by extrinsic order[2] in the first

---

[2] Independent justification for the ordering of the unbounded movement rules before Extraposition-from-NP is supplied by the paradigm below:

(i) a. You believe the girl who was sick to have left.
b. *Which girl do you believe who was sick to have left?

(ii) a. You saw a girl who was wearing a red hat.
b. Who did you see who was wearing a red hat?
c. Who who was wearing a red hat did you see?

As (i) shows, the head of a relative clause cannot be moved by Question-Movement. But precisely such a movement seems to have derived (iib) from (iia) without causing unacceptability. The acceptability of (iib) can be explained if Question-Movement derives (iic) from (iia) and Extraposition-from-NP then derives (iib) from (iic). The ordering Extraposition-from-NP--Question-Movement could not derive (iib) from (iia), for the former rule is inapplicable to (iia), unless we allow vacuous Extraposition-from-NP, a move which has generally been deplored by transformational grammarians. In addition, Ross (1967) shows there are reasons for disallowing the questioning of a relative clause head after Extraposition-from-NP, as suggested by the ungrammaticality of (iiib), derived from (iiia) with the ordering Extraposition-from-NP--Question-Movement; in contrast, (iiic), derived from (iiia) with the opposite ordering, is more acceptable.

(iii) a. Sam didn't pick up those packages which are to be mailed tomorrow until nine o'clock.
b. *Which packages didn't Sam pick up which are to be mailed tomorrow until nine o'clock?
c. Which packages didn't Sam pick up until nine o'clock which are to be mailed tomorrow?

---

and by a global rule like (1.86) in the latter two.

(4.8) a. A boy who kissed Mary just walked out.
b. *Who did a boy who kissed just walk out?
c. A boy just walked out who kissed Mary.
d. *Who did a boy just walk out who kissed?

(4.9) a. I think that what John wants it to kill the girl.
b. *The girl who I think that what John wants is to kill is pretty.
c. It's John's courting that girl that I strongly condemn.
d. *The girl who it's John's courting that I strongly condemn is pretty.

4.1.1.3. The third problem to be discussed concerns the paradigm in (1.90), where the greater acceptability of (1.90b) as against (1.90a) requires a global rule like (1.91).

(1.90) a. *It's Mary (who) I discussed the claim that Bill likes.
b. ?It's Mary (who) I made the claim that Bill likes.
c. It's Mary (who) I claimed that Bill likes.

(1.91) Violations of island constraints are considerably reduced in severity if the affected island was not an island in underlying representation.

In 1.2.6, I pointed out that (1.91), rather than extrinsic ordering, is needed to account for (1.90), because the configuration make the claim is needed in the cycle for cyclical rules like Passive and Raising to yield (4.10a), or for Tough-Movement to yield (4.10b).

(4.10) a. The claim that Bill likes Mary is believed to have been made by Jo.
b. The claim that John likes Mary is hard for me to make.

What needs to be explained is why the postcyclic unbounded rules which move elements out of a complex NP should be sensitive to whether the latter is an underlying or derived island.

4.1.2. The theory of perceptual strategies was described in Chapter Two. One point, however, bears repetition: the real-time operations by which sentences are decoded are not, in all likelihood, mechanical replicae of the rules of a generative grammar applying in reverse order. One argument often given in support of this position is that the perceptual and derivational complexities of sentences do not always coincide. An even more compelling argument against the claim that perception 'undoes' derivations is based on the observed limitations of echoic memory. Thus, a great deal of work in cognitive psychology suggests that auditory memory is of at least two kinds: echoic and verbal (the terminology is Neisser's (1967)). Echoic memory is assumed to be a short-term store for unsegmented signals which have undergone only global preprocessing operations, while verbal memory is regarded as a more permanent store for coded material obtained through a synthesis of the echo by the application of perceptual strategies. There have been various attempts to measure the duration of echoic memory; thus, Guttman and Julesz (1963) estimate the duration of echoic memory as approximately one second long, Pollack (1959) puts it at about four seconds, while Ericksen and Johnson (1964) allow as much as ten seconds. The range of variability of the various measurements seems to pose a problem, but Neisser (1967, p. 205) points out that the time values obtained in the three sets of experiments were inversely proportional to the difficulty of the tasks that the subjects were called upon to perform. If we regard speech perception as a task of intermediate difficulty, a fair estimate of echoic memory would place its useful duration at about four seconds. Consequently, a theory of perception which claims that transformations are undone anticyclically, and which therefore requires that no processing take place until the whole signal has been stored for the highest cycle processes to be activated, falsely predicts not only that nothing is understood until the end of a sentence is heard, but also that only that portion of a complex sentence which falls within the final four-sentence span can be understood. For obvious reasons, a grammar which incorporates a pre- and/or a post-cycle in addition to the cycle is even more implausible as a real-time model.

It seems more plausible to assume that there are units of syntactic perception (which, according to Lehiste (in press) are determined by suprasegmental structure) to which perceptual strategies are applied in the order in which the units are presented. In addition to these 'local' strategies, we must assume 'integrative' strategies (which very remotely resemble the generalized transformations of early transformational grammar) which utilize suprasegmental cues, 'function words', discontinuous constituents, etc., in order to synthesize the hierarchical relations between the units of syntactic perception into a coherent formal object. Thus, the strategies are presumed to make use of every cue in order to emit a fairly narrow range of hypotheses concerning what may or must follow, and to strike out hypotheses incompatible with later in-coming cues, until the desired reading of the sentence has been zeroed in on (or readings, if the sentence is linguistically and/or contextually ambiguous); of course, the strategies are heuristic procedures, and, unlike the analytic procedures of American structuralism, offer no guarantee of success.

Concerning the order of application of perceptual strategies, there have been some experiments suggesting that some strategies apply simultaneously while others apply sequentially. Thus, Mehler and Carey (1968) presented subjects with appropriate and inappropriate pictures followed by a sentence containing either progressive or participial constructions like (4.11a) or (4.11b) respectively (it was assumed that the latter was syntactically more complex than the former).

(4.11) a. They are fixing benches.
b. They are performing monkeys.

The response times were fastest when the progressive construction was used and the picture was appropriate, but delayed by the same amount otherwise. Thus, it seemed to make no difference whether both the syntax and the semantics, or only one of them, were complex, which suggests that at least certain aspects of meaning and syntactic structure are processed simultaneously. Concerning the evidence for sequential application of strategies, Bever (1971) reports on a number of click-localization experiments which suggest that subjects are immediately aware of clausal[3] boundaries, but not of boundaries to wait to the end of the sentences before reporting the location of clicks; she repeated the experiment asking subjects to react as soon as the click was heard, and got different results.

---

[3] Lehiste (in press) points out that the recognition of clausal boundaries is probably itself not an immediate process, and occurs prior to the recognition of suprasegmental structure. She suggests that Bever's results were due to the fact that he asked his subjects

It should also be pointed out that Bever also claims that the primary units of syntactic perception are *deep*, not *surface*, clausal boundaries. He bases his claim on the observation that clicks occurring in the middle of *the troops* tended to migrate to the right in (i) and to the left in (ii); thus, subjects seem to be sensitive to deep structure boundaries.

(i) The general ordered the troops to fight.
(ii) The general desired the troops to fight.

Notice, however, that clicks cannot be attracted to deep clausal boundaries unless the elements separated by clausal boundary in deep structure are still adjacent in surface structure. Thus, it is not clear what the click should be attracted to in (iii) in order to support Bever's hypothesis.

(iii) John definitely seems to be stupid.

In (iii), the deep boundary is between *seem* and *John*. Bever's hypothesis becomes totally implausible if a deep structure allowing lexical decomposition is considered, for it would imply that if we decompose *no man came* as (NOT [SOME man came]), the boundary between NOT and SOME is a primary syntactic percept.

---

between non-sentential constituents; however, if subjects are asked to wait several seconds before reporting the location of the click, non-sentential constituent boundaries do appear to cause response bias. This suggests that some structural features are processed earlier than others.

Concerning the strategies which apply sequentially, we must recognize at least two kinds of ordering, O_a and O_b. O_a is an exact replica of intrinsic grammatical ordering; that is, if some strategy A creates the very input to some strategy B, it follows that A must apply before B,. As a hypothetical illustration, consider the following possible way in which the processing of (4.12) may proceed.

(4.12) *Who* do you think John fears that Mary dislikes?

If the primary units of syntactic perception are suprasegmental rather than structural (as proposed by Lehiste), there must be strategies which recognize clauses and segment them into predicates and arguments. One way in which this may be done is through the identification of verbs and the retrieval of the syntacto-semantic information associated with them (from some long-term store corresponding to the lexicon); by recognizing the number and kind of potential arguments that a given verb requires, it is possible to emit tentative hypotheses about which arguments go with which verbs. The underscored word in (4.12) is easily recognizable as a highly probable member of a discontinuity, and as the analysis by synthesis of the incoming signal proceeds, the wh-word is tested as a possible argument of the various verbs encountered. Specifically, who is not considered as a possible object of think, as the latter requires abstract objects, nor of fear, for although this verb may take animate objects, it can take only one object, and is already followed by one; therefore, who can only be an argument of dislikes. If this oversimplified account is in any way correct, the strategy which attempts to match who with some verb can only apply after information as to the selectional restrictions of that verb has been retrieved.

O_b depends on the number of cues required for the triggering of specific strategies, on the moment in time when those cues become available, and on the complexity of the task which the strategies in question are called upon to perform. I believe that O_b can provide an explanation for the problems raised in 4.1.1.1 and 4.1.1.2. Of course, there is no reason why specific strategies should not operate simultaneously, if the pertinent cues are simultaneously provided,[4]

which is, I suggest, the explanation for the problem raised in 4.1.1.3.

4.1.3. In this section, I propose solutions to the problems in 4.1.1.

4.1.3.1. Before considering a solution to 4.1.1.1, it is important to recognize that (4.1) is probably a perceptual, rather than a formal, constraint. The point is that it is difficult to identify a constituent which forms a perceptual unit with the flanking material when the category to which that constituent belongs is hard to identify. Thus, the internal constituents in (4.2a) and (4.2b) are arguments of a higher predicate and therefore NP's; however, they are too sentence-like to be readily recognizable as NP's. On the other hand, the internal constituents in (4.2c) and (4.2d) are sufficiently noun-like (they begin with genitives, i.e., determiners) to be quickly identified as such. With respect to (4.3), where the underscored constituent is followed by some suprasegmental boundary, we are faced with a clause internal to the containing clause, but not to the containing perceptual unit (recall Lehiste's hypothesis that perceptual units are determined by suprasegmental structure), and its recognition as object of explain is thereby facilitated. It appears that the notion 'internal NP-over-S' must be reformulated in terms of containing perceptual units, rather than clauses, which is not at all surprising if (4.1) is a perceptual, rather than a purely formal, constraint.

If (4.1) represents a limitation of the strategy which identifies arguments of higher predicates, then the contrast in acceptability

---

[4] It is not clear whether simultaneous application of decoding operations involves parallel independent processing or very rapid shifting back and forth from one task to another. If we accept Neisser's (1967) views that synthesis, as opposed to global preprocessing, is strictly sequential, we are forced to entertain only the latter hypothesis.

between (4.4a) and (4.4b) and (4.7b) is rather easy to explain: the heads of free relatives or embedded interrogatives are signaled by the characteristic *wh*-pronoun, and can therefore be reconstructed *before* the clause is assigned a role in the containing sentence; in other words, the strategy which reconstructs the heads of *wh*-clauses can operate *before* the strategy which segments out the arguments of a higher predicate, in which case the latter is no longer blocked by (4.1), for the internal clause is no longer *exhaustively* dominated by NP. In contrast, a mere inspection of the underscored constituent in (4.7b) will not tell us whether this constituent does or doesn't have a deleted head; it will be necessary to wait for the higher predicate, and even then the reconstruction of *the fact* cannot employ overt clues, like the *wh* feature, but must base itself on the meaning of *realize*, from which it may be inferred that the truth of the underscored constituent is presupposed, and that it consequently modifies a head *the fact*. Thus, the availability of cues as well as the relative complexity of the task that strategies have to perform is consistent with the ordering in (4.13), which provides a straightforward explanation of the contrast between (4.4) and (4.7b).

(4.13) (i) Reconstruction of the heads of *wh*-clauses;
(ii) Role assignment to constituents within the containing sentence (constrained by (4.1));
(iii) Reconstruction of the heads of factive complements.

(4.13) is a little too strong in asserting that step (iii) does take place; in fact, the data we have considered allow only the weaker statement that step (iii) (*if it takes place at all*) must follow both steps (i) and (ii). This weaker statement is in order since there is no evidence that *the fact* is reconstructed at all in the decoding of (4.7b).

In R. Lakoff (1972), it is argued that there exist certain difficulties with the Kiparskys' analysis of factives, and it is suggested that their behavior might be more adequately accounted for by invoking their semantic properties (in a way unknown at present) than by postulating a head NP *the fact*. Notice, however, that even if Lakoff's arguments are accepted and the Kiparskys' analysis is rejected, it does *not* follow from this that hearers *necessarily* do not reconstruct a head for the complement of *realized* in interpreting (4.7b). To see this, consider embedded interrogative clauses like *I know who left*, which have sometimes been analyzed as *I know the answer to the question: who left*. The status of the head *the answer to the question* is at least as questionable syntactically as that of the head *the fact* which the Kiparskys propose. For example, adverbs modifying interrogative clauses can sometimes be preposed beyond the matrix verb, as in *in case you have a minute, I wonder whether I could talk to you* (cf. **in case you have a minute, I wonder about the answer to the question whether I could talk to you*); also, some interrogatives are not islands for the purpose of unbounded rules, as shown by *it's Mary I wonder why Bill left* (cf. **it's Mary I wonder about the answer to the question why Bill left*). Now, the fact that interrogatives do not seem to modify a Nucleus in some cases does not mean that such a Nucleus is never reconstructed, for is why he left a real mystery? or it's who went where that I am puzzled by are just as good as is what you found on the table the tissue? Consequently, it appears that hearers do reconstruct heads for interrogatives, although the syntactic justification for such heads is just as shaky as that for the heads of factives.

If one wanted to deny that internal interrogatives are reconstructed with heads, one would have to deny that the exhaustive domination of the node S by the node NP has anything to do with the unacceptability of (4.2a) or (4.2b). While such a step is possible in principle, I propose that it not be taken until a viable alternative is proposed.

4.1.3.2. We now turn to problem 4.1.1.2, which consisted of the difficulty of stating the ungrammaticality of (4.8d) and (4.9b) as the same phenomenon. Due to specific properties of the formal descriptive model, the lower level discontinuity has to be created before the higher level one in (4.8d), while the opposite ordering obtained in (4.9b). The two cases will, however, be reduced to the same case if it can be shown that when the integration of the two discontinuities is attempted in real time, the existence of the higher level discontinuity is recognized before the integration of the lower level one is carried out. That is, when the interrogative or relative pronoun is matched with the pseudo-cleft constituent or the extraposed clause respectively, the perceptual device must have been already apprised of the fact that these larger constituents have been torn off complex NP's. It seems to me that this condition is guaranteed by the order of presentation of the various clues. The pseudo-cleft construction is a highly characteristic one, consisting of a free or full relative clause, a copula, and a pseudo-cleft constituent, in that order; therefore, by the time the pseudo-cleft constituent is presented, the perceptual device will have been given more than enough cues to recognize that this constituent enters a discontinuous relation with the subject relative clause[5] (sentences

---

[5] Of course, a sequence of relative clause followed by copula need not be continued with a pseudo-cleft constituent, as (i) shows.

(i) What John ate was wonderful.

However, one of the things that speakers know about pseudo-cleft constructions is that the pseudo-cleft constituent enters into selectional restrictions with some verb in the relative clause. As there is no *John ate wonderful, a hearer will easily recognize that (i) is not a pseudo-cleft construction.

As pointed out in Bach (1969), there exist ambiguous constructions like what Descartes discovered was a proof of his existence. In such cases, we predict that the ambiguity is resolved when an element of the predicate has been chopped; this is confirmed by (ii), where the a-sentence is ambiguous, but where the b-one is not, as like (4.9d) can easily be accounted for in the same way, as the cleft construction is also highly characteristic). With respect to extraposed clauses as in (4.8d) or (4.14),

(4.14) Who do you believe $\left\{\begin{matrix}\text{it}\\ \text{*the claim}\end{matrix}\right\}$ is odd that Bill kissed?

they are recognizable as relatives or complement clauses immediately, for they _must_ be introduced by complementizers; indeed, as (4.15) and (4.16) show, the complementizer can be deleted when the clause is in construction with its head noun, but not if it is extraposed.

(4.15) a. A boy you met recently has just arrived.
b. *A boy has just arrived you met recently.

(4.16) a. The rumor John is sick is unconfirmed.
b. *The rumor is unconfirmed John is sick.

A clause in sentence-final position can be either an argument of the sentence predicate or a dependent of some head noun; by the time the extraposed clause is presented, the remainder of the sentence has already been heard, and it becomes possible to see whether the clause is more appropriately matched with a preceding NP or with the predicate.

If the clause is matched with an earlier NP, any discontinuity involving this clause and some external element will be recognized as a violation of the CNPC.

4.1.3.3. Concerning the problem in 4.1.1.3, namely, the distinction in acceptability between (1.90a) and (1.90b) can be accounted for quite straightforwardly in terms of the Mehler and Carey experiment mentioned in 4.1.2. Thus, assuming that speakers have an internalized version of the lexicon, in which lexical items are characterized phonologically, syntactically and semantically (as I suggested in the discussion of (4.12)), the recognition of a lexical item should involve the retrieval of its characterizing information, its syntactic and semantic properties thus becoming available simultaneously. Therefore, at the point at which a hearer attempts to integrate the discontinuity in (1.90b), he has already recognized that _the claim_ looks like the head of a complex NP, but is not one, semantically. Quite likely, _make the claim_ is stored as one lexical item, in fact, a sentential idiom, and it is a general fact about idioms that (at least part of) their apparent structure is not relevant for interpretation. As we shall see in 4.2.5 structural

---

the non-pseudo-cleft reading alone is grammatical.

(ii) a. What I saw was a representation of that building.
b. That building, which what I saw was a representation of, has collapsed.

configurations not relevant to semantic interpretation cannot function as islands even when they are represented as islands in underlying representation.

The somewhat more awkward character of (1.90b) as compared with (1.90c) may be due to the conflict between what the phrase *make the claim that...* seems to be and what it really is. Notice that the simultaneity of the processing of syntactic and semantic cues is important; for, if the latter were processed significantly later than the former (as one would expect, if perception undid the derivations of, say, an *Aspects*-model grammar), (1.90b) would very probably be perceived as ungrammatical in the initial processing stages, and further processing would be discontinued, thus never giving the semantics a chance to right the situation (cf. (1.90b), which is immediately interpretable, with Bever's *the horse raced past the barn fell*, which might not be if *raced* is interpreted as a preterite).

4.1.4. I have tried to show in section 4.1 that the ordering of strategies has significant effects on the acceptability of sentences.

Thus, I argued in 4.1.3.3 that the simultaneity (or near-simultaneity) of syntactic and semantic processing can account for the almost complete acceptability of (1.90b).

In 4.1.3.1, I attempted to show that the ordering proposed in (4.13) can provide an explanation for the contrast in acceptability between (4.4a) and (4.4b) and (4.7b). That is, in order to account for the facts, the heads of the internal clauses *must* be recognized before the clauses are assigned a role in (4.4a) and (4.4b), but after that in (4.7b). The unacceptability of (4.7b) is particularly important, for it shows that (4.13ii) fails if the head of the internal clause is recognized 'too late'.

In 4.1.3.2, I showed that the two sentences at issue, namely (4.8d) and (4.9b) can receive the same explanation, since in both cases the higher level discontinuity is recognized before the lower level one is resolved. However, in order to show that the order in which the discontinuities are integrated is *significant* rather than *accidental*, it would be necessary to find a *grammatical* sentence in which the existence of the higher level discontinuity would not be recognizable when the lower level one is resolved. Constructions satisfying this requirement may be hard to come by. I have, nevertheless, been able to construct an approximately adequate paradigm in Japanese, which fortunately has both leftward and rightward unbounded processes capable of violating the CNPC; such processes are Topicalization and Relativization respectively. In Ross (1967), it is pointed out that complement clauses modifying 'lexical' head nouns like *syutyoo* (the claim) are impervious to Relativization, while clauses modifying the dummy *koto* (thing) are not. As the head nouns of such constructions invariably appear on the right of the modifying clauses, one may hope that Topicalization would be free to move elements out of complement clauses, since at the stage at which the discontinuity is integrated, there would be no way of knowing whether the head of the complement clause is a lexical element or a dummy. In other words, there should be no more difference in acceptability between (4.17c) and (4.17d) than there is between (4.7a) and (4.7b), while there should be a difference between (4.17b) and (4.17d).

(4.17) a. Otto ga kabutte ita koto e watakusi ga
Otto wearing was thing I
sinzita boosi wa akai.
believed hat red
'The hat which I believed that Otto was wearing is red.'
b.(*)Otto ga kabutte ita to iu syutyoo o watakusi
Otto wearing was that say claim I
ga sinzita boosi wa akai.
believed hat red
'The hat which I believed the claim that Otto was wearing is red.'
c. Otto wa, watakusi wa kono boosi o kabutte
Otto I this hat wearing
ita koto o sinzita.
was thing believed
'Otto, I believed was wearing this hat.'
d.(*)Otto wa, watakusi wa kono boosi o kabutte
Otto I this hat wearing
ita to iu syutyoo o sinzita.
was that say claim believed
'Otto, I believed the claim was wearing this hat.'

However, the few Japanese informants I have been able to consult have assured me that (4.17c) and (4.17d) differ in acceptability, while (4.17b) and (4.17d) do not.[6] While (4.17d) does not offer a perfect test case, since the head and the complement clause are adjacent (i.e., there is no higher level discontinuity), its unacceptability is nevertheless sufficiently significant to suggest that the fact that (4.8d) and (4.9b) could receive the same explanation was something in the nature of a lucky accident.

The lack of a distinction in acceptability does not, however, run counter to the explanation I offered in 4.1.3.2. Thus, the fact that we find no distinction in acceptability between (4.7a) and (4.7b), but we do find one between (4.17c) and (4.17d), is due to the fact that there is a distinction between the NP-over-S Constraint and the CNPC, the former being a perceptual, and the latter a grammatical principle (in fact, a grammatized one, as I will argue in 4.3). Put differently, the latter, but not the former, is a part of the competence of speakers, and while we would expect a perceptual difficulty to arise only when the string to be processed is complex in the appropriate way, there is no reason why such an expectation should be entertained with respect to a grammatical (or grammatized) constraint. Specifically, even though a hearer presented with (4.17d) would in all probability not know, at the crucial moment, that the complement clause with which the topicalized constituent is integrated is part of a complex NP, he would still know that violations of the CNPC are unacceptable and would therefore strike out the possibility that the clause in question might modify a lexical head noun. The discovery that syutyoo, rather than koto, is the head of the complement clause conflicts with the initial analysis of that clause, and unacceptability arises at that stage.

---

[6] The star in parentheses in (4.17b) and (4.17d) indicates that the corresponding sentences are unacceptable in some dialects, but not in all dialects.

It is worthwhile to pause for a while and consider how the above account, if correct, fits within the perceptual framework we have been using so far.

Clearly, the difficulty arising from the categorization of the complement clause modifying syutyoo in (4.17d) as both a Satellite and a non-Satellite is a typical instance of perceptual conflict. In addition, it is interesting to note that the initial discarding of the possibility that the complement clause be a Satellite bears certain similarities to erroneous closure; thus, in both situations, an a priori theoretically possible analysis of some substring is ruled out 'too soon'. An even more interesting case which may be regarded as involving 'erroneous partial closure' will be discussed in 4.2, in relation to the relative complexity of low predictability. In essence, it will be argued that the lesser complexity of predictable structures leads to the hypothesis that they are most likely to occur than non-predictable ones; consequently, the possibility of finding some desired constituent in an unlikely place is 'partially ruled out', as it were, and this, it is claimed, is partially responsible for the SSC.

To return to the problem discussed before, an English paradigm which demonstrates the same point as (4.17) is perhaps (4.18); (4.18) is not an ideal test case either, since informants are far from unanimous in accepting its acceptability ratings.

(4.19) a. John referred to the roof of the house in his discussion with Mary.
b. I've just seen two birds on the roof of the house.
c. It's the house which John referred to the roof of in his discussion with Mary.
d.*?It's the house which I've just seen two birds on the roof of.
e. The roof of the house, John referred to in his discussion with Mary.
f. The roof of the house, I've just seen two birds on.
g. ?It's the house which the roof of, John referred to in his discussion with Mary.
h. *It's the house which the roof of, I've just seen two birds on.

The crucial constituent is *the roof of the house*, which forms an adverbial Satellite in b, d, f, h, but not in a , c, e, g. The c and d pair shows that the maiming of that constituent is considerably worse in the adverbial case; the e and f pair shows that the constituent in question can be topicalized in either case; the g and h pair, in which the constituent is both topicalized and maimed, shows that the overall result is considerably worse when the constituent at issue *originated* as an adverbial.

Clearly, at the stage at which the topicalized element is encountered in the g or h sentence, there is no way of telling whether it is a Satellite or not. But the h sentence is much worse than the g one, and this fact again suggests that language users call upon their knowledge of the Subservient Satellite Constraint to mark the topicalized constituent as a non-Satellite. As in the Japanese case, this initial analysis is incompatible with the remainder of the string in (4.18h), and the sentence is therefore unacceptable.

4.2. This section is essentially devoted to the mitigating effect of predictability (and the aggravating effect of unpredictability) on the complexity of interrupted behavior.

I begin by reconsidering one parameter which Bever regarded as significantly determining the complexity of discontinuities, namely, structural complexity (see (2.81)), and which, to the best of my knowledge, has never been satisfactorily defined. Ross (1967) discusses the notion of 'complex NP' in relation to the acceptability of sentences which have undergone Particle-Movement, and proposes that 'an NP is complex if it dominates the node S'. On this basis, he accounts for the greater unacceptability of (4.19c) as compared with (4.19a) or (4.19b).

(4.19) a. John called the unusually pretty girl up.
b. John called the girl from San Francisco up.
c. ?John called the girl who was from San Francisco up.

Ross also points out that structural complexity is probably not a simple function of length, as he finds (4.20a) more acceptable than (4.20b).

(4.20) a. I called almost all of the men from Boston up.
b. I called the man you met up.

It appears therefore that, with respect to Ross' dialect (for not all speakers share his judgments about (4.20)), structural complexity depends, to some extent, *on the occurrence of the node S between discontinuous components*. However, it appears that even though most speakers find (4.19a) better than (4.19c), some speakers find (4.20b) as acceptable, or even more acceptable, than (4.20a). This suggests that the length, as well as the number of nodes (i.e., constituents and subconstituents), of the intervening material plays some role in determining its structural complexity, and the dialectal differences are presumably to be ascribed to the different 'weights' given to each contributing factor in specific idiolects. The situation is in fact even more complicated, for certain speakers find (4.19b) intermediate in acceptability between (4.19a) and (4.19c), even though the a and the b-sentences have roughly the same length and number of intervening nodes. Several explanations for this distinction could probably be found, one of which could be the higher degree of (perceptual) recoverability of the pruned S-node over from San Francisco as compared with the pruned S-node over unusually pretty.

Clearly, a great deal more research must be undertaken before a satisfactory explication of the notion of structural complexity is arrived at. In the remainder of section 4.2, I will consider several cases which suggest that a principle like (4.21) must be incorporated into any effort to further explicate structural complexity.

(4.21) The structural complexity of a string is inversely proportional to its structural predictability.

4.2.1. Contrast (4.22a) and (4.22b).

(4.22) a. John called the not very well liked but quite pretty girl on the next block where Jack had lived for six years up.

b. Who do you think Mary told her friends that she saw at the party?

Most speakers find (4.22b) considerably more acceptable than (4.22a). If one of the main factors which make discontinuities complex is the number of intervening sentence nodes, then it is not clear why sentences like (4.22b) should be better than (4.22a), since there are more S-nodes between who and at the party in the surface structure of (4.22b) than between called and up in the surface structure of (4.22a). I believe the reason is that the structure of the intervening material in (4.22b) can be essentially described with a simple recursive procedure, while the intervening material in (4.22a) is not amenable to such treatment. In other words, the intervening material in (4.22b) can be generated by recursively applying the set of rules in (4.23), while many more rules need to be specified for the generation of the intervening material in (4.22a).

(4.23) (i) S → NP VP
(ii) VP → V NP
(iii) NP → S

Thus, to say that the pertinent material in (4.22b) is more predictable is to say that there are fewer unrelated structural relations between its subconstituents which a language perceiver must discover while holding the first member of the discontinuity in memory. Consequently, the perceiver needs to carry out less processing on that material, and the burden on his immediate memory is alleviated.

But, one may ask at this point, how does the language perceiver know in advance, at any point, that the incoming material will not necessitate further processing, and that it can be taken care of, by and large, by reapplying an already discovered procedure some number of times? The answer is that a language user, who has had a certain amount of experience with language, has already extracted the generalization that some structures are recursively predictable while others are not. Thus, the language user knows (by induction over experience, or with the help of innate mechanisms, or both) that structures like (4.22b) where the resolution of the discontinuity is to be sought somewhere 'down the chain of command' by a procedure roughly like (4.23) are comparatively non-complex structures, and therefore possible ones. Therefore a hearer will project a recursive structure of the appropriate form, and, when scanning the linear signal, will experience much less difficulty if the second member of the discontinuity occurs where predicted by the recursive procedure.

It is rather obvious that additional processing does affect acceptability. Thus, (4.24) or (4.25), in which the intervening material requires not only the recursive application of (4.23) but also some additional processing, are less acceptable than (4.22b).

(4.24) It's John who I believe Joe, Dick and Mary told the two guys who were in love with Jill and Donna respectively that Bob hates most.

(4.25) Who do you think my girl friend from Chicago suggested to the boy who lives across the street that Jill loves best?

Let us consider at this point Ross' HIC (see Chapter One) which claims that a chopped element must command its place of departure (in fact, this constraint applies not only to chopping, but to all the situations in which the stronger island constraints apply, i.e., feature-changing, deletion under obligatory coreferentiality, semantic neutralization, etc.). Ross hypothesized that the HIC is a universal fact about human languages, and, as far as I know, no counterexamples have been discovered. I believe that the HIC can be at least in part explained by (4.21), since (as I shall argue in detail in section 4.4) the situations in which the various island constraints--including the HIC--hold, are precisely the situations in which a structural discontinuity exists. Notice then that the HIC reduces the complexity of discontinuities in at least two ways: (i) by increasing the predictability of the interdiscontinuous material through a recursive procedure, i.e., (4.23), and (ii) by making it easier to predict the location of the second member of the discontinuity, as the latter needs to be sought only down the chain of command, rather than anywhere in the tree.

In the remaining subsections of 4.2, I will discuss several phenomena involving discontinuities in which the following situation obtains:

(4.26) Acceptability increases in proportion to the predictability of (i) the structure of the interdiscontinuous material, and/or (ii) the location of the second member of the discontinuity.

4.2.2. In 2.1.3, I proposed the perceptual principle (2.83).

> (2.83) Two interrupted sentences are perceptually complex if one of them is at least in part internal to the other.

One of the cases in which (2.83) takes effect concerns the RRC", where there exist two discontinuities, such that the second is created by the resolution of the first. The difficulty is to resolve both discontinuities properly when the location and/or structure of the second member of each discontinuity is not easily predictable, as in (4.27). In note 19 of Chapter Two, I argued that sentences like (4.28a) are much more acceptable than (4.27) because the former does not satisfy (2.83), the interrupted sentences being all resolved simultaneously and none being therefore internal to any of the others.

(4.27) ??What I will tell the jury that John thinks about tomorrow is his luscious mistress.

(4.28) a. Mary cooked, John found, and Bill ate, the cake.
b. John bought, Mary cooked, and Bill drank, a bicycle, a cake, and a bottle of champaign respectively.

The greater acceptability of (4.28a) with respect to (4.28b) can be additionally illuminated by (4.26). Indeed, in (4.28a), as soon as the hearer has realized that *Mary cooked* is a maimed sentence, and that what follows is coordinated to it, he will be able to predict that the following coordinate terms will all be maimed in 'parallel fashion' (in the sense of note 14, Chapter Three), whatever their number. Also, as claimed by (4.27i), the number of interrupted sentences will not seriously increase the complexity of the inter-discontinuous material, since the latter is essentially predictable by the recursive application of the rule $X \rightarrow X\ (\left\{\begin{matrix}\text{and}\\ \text{or}\end{matrix}\right\})\ X$. With respect to the requirement made by (4.26ii), the knowledge language users have about the scope and nature of Coordination Reduction makes it possible to predict that the resolution of the discontinuities will follow immediately upon the presentation of the last maimed coordinate term.

The most interesting case for our purposes is (4.28b), since *it violates (2.83)*, the interrupted sequences being separately continued, and nonetheless it is much more acceptable than (4.27). Notice that in (4.28b) the resolution of the various discontinuities is just as predictable as in (4.28a). Indeed, a sentence like (4.28a) arises when the elements moved by Coordination-Reduction are all identical, and a sentence like (4.28b) arises when the moved elements are not all identical; in the latter case, we canppredict that: (a) the resolution of the discontinuities will be a constituent *of the same kind* as the constituent formed by the interrupted sequences, namely, a coordination; (b) the two coordinations will have *the same* number of terms; (c) the discontinuities will be resolved in the order in which they were presented, i.e., the nth term of the first coordination will be resolved by the nth term of the second coordination.

Thus, we can see that considerations of predictability can improve the acceptability of sequences which one would otherwise expect to be unacceptable (in the case of (4.28b), in view of (2.83)).

4.2.3. The third phenomenon I shall consider with respect to predictability is the SSC (Sentential Subject Constraint). Ross (1967, section 4.4.2) points out that the unacceptability of sentences in which the SSC has been violated is due, to a large extent, to the fact that the NP-over-S Constraint is also necessarily violated by the movement of the chopped constituent to the left of the headless sentential subject (he gives credit to G. Lakoff for this observation). However, Ross also points out that the SSC is still necessary because violations of the SSC, added to violations of the NP-over-S Constraint create slightly greater unacceptability than violations of the latter alone. I reproduce two of the examples with which he illustrates his claim below; in (4.29b) and (4.30a), the NP-over-S Constraint is violated, in (4.29c) and (4.30b), both the NP-over-S Constraint and the SSC are violated.

(4.29) a. That I brought this hat seemed strange to the nurse.
b.??The nurse who that I brought this hat seemed strange to was as dumb as a post.
c. *The hat which that I brought seemed strange to the nurse was a fedora.

(4.30) a.??I deny that that McIntyre has some money is certain.
b. *I deny that that McIntyre has any money is certain.

A second reason mentioned by Ross for believing that combined violations of the SSC and the NP-over-S Constraint are worse than violations of the latter alone has to do with the Nominality-Sentencehood scale (see Appendix Two). Thus, the NP-over-S Constraint is applicable when the internal constituent is a that- or for-to-clause, but not a gerundial, nominalization, or anything even more nominal; the SSC, on the other hand, is applicable when the maimed subject is a that-clause, a for-to clause, or a gerundial. The contrast in applicability to gerundials of the two constraints is shown in (4.31a) and (4.31b) respectively.

(4.31) a. Is playing that piano all right with you?
b. *That piano, which John's playing is offensive to me, is a Steinway grand.

Thus, it seems that the applicability of the SSC on top of the NP-over-S Constraint extends the domain of the latter one further step towards the nominality pole.

The reduction in acceptability which (4.29)-(4.31) indicate should be associated with the SSC, is, I suggest, due to the fact that in structures in which the SSC has been violated, a simple recursive search procedure for the second member of the discontinuity becomes inapplicable: the result is greater psychological complexity, as predicted by (4.26ii). Indeed, notice that the search for the slot of a constituent which has been chopped out of an object clause can proceed as follows: (a) search for the verb of the first clause which occurs in the inter-discontinuous material (this has to be done in any case, since the segmentation of sentences is probably carried out, to a large extent, by comparing the number of expected arguments of the verb with what is actually present in the string); (b) move to the right of the verb, and see whether the chopped element can be an argument of that verb; (c) if the answer to (b) is no, the verb has a clausal object; reapply step (a) to this clause, and continue until (b) succeeds. The procedure outlined above moves to the right at every step; however, if the chopped element belongs in a subject, the procedure must look outside the recursively predictable areas. This step causes some perceptual complexity.

An independent piece of evidence which supports my account of the SSC is provided by the maimability of topicalized constituents, which exhibit the same paradigm as subjects with respect to the nominality scale, as can be seen by comparing (4.32) and (4.33).

(4.32) a. *It's John who that you killed is clear.
b. *It's John who for you to meet may be dangerous.
c. *It's John who meeting may prove dangerous.
d. It's John who the shooting of has become imperative.
e. It's John who a picture of is in my pocket.

(4.33) a. *It's John who that you killed yesterday the FBI finds incredible.
b. *It's John who for you to meet at noon the FBI finds desirable.
c. *It's John who meeting at noon the FBI regards as dangerous.
d. (?)It's John who the shooting of in cold blood the FBI has been too ready to condone.
e. (?)It's John who a picture of taken on D-day the FBI has been trying to obtain.

Not all informants tolerate topicalization in embedded clauses, but those who do feel that the d and e sentences are considerably better than the remaining ones in (4.33). Thus, for the speakers who agree with the judgments in (4.32) and (4.33), an explanation is available on the grounds that the recursively predictive procedure I proposed above is violated in the same way by the maiming of subjects or of topicalized constituents. Notice that these two phenomena are not statable as one constraint in Ross' framework, since some mark over topicalized constituents (such as ADV, or TOP) is probably needed in order to distinguish formally between topicalized elements and subjects; consequently, the explication of (4.33) along Ross' lines would, in all probability, require an ad hoc constraint in addition to the SSC.

An interesting prediction that my account of the SSC makes is that the SSC is operative only in surface SVO languages, as it is only in such languages that arguments may be positioned on both sides of a verb. The languages in which Ross pointed out that the SSC holds are indeed SVO ones, while the only language he mentions in which the SSC does not hold is Japanese, an SOV language. The failure of the SSC to hold in Japanese is illustrated in (4.34).

(4.34) John-ga suki ta to iu koto ga hakkiri shite iru
John likes obvious
so no uchi wa kiree da.
house pretty.
*'The house which that John likes is obvious is pretty.'

Of course, no hypothesis can be regarded as confirmed on the basis of one language, but I have found two more languages in which the SSC fails to hold, and they are both SOV.

Thus, Hankamer (1971) cites the following example from Turkish:

(4.35) Mehmedin yiyecegi şüpheli olan meyva.
Mehmet's going to eat doubtful being fruit
*'The fruit which that Mehmet will eat is doubtful.'

The well-formedness of (3.35) is significant, for, as Hankamer points out, the Turkish rule of Relativization--which always operates by deletion, like the Japanese one--is subject to the CNPC.

The second SOV language which disobeys the SSC is Hindi. Hindi Relativization operates by movement toward the relative clause head, which usually precedes the clause.[7]

---

[7] In fact Hindi supports the claim made in Drachman (1970) that Relativization proceeds by copying and deletion. Thus, a pro-form may optionally be left behind, as (i) shows.

(i) vo admi Ĵiske bareme betaiahe ki $\begin{Bmatrix} \text{vo} \\ \text{—} \end{Bmatrix}$ bimarke
the man whom about I told you the (he) sick
Ĵacukahe.
has left
'The man about whom I told you is sick has left.'

In addition, the relative clause and its head may be permuted, in which case, a copy must appear in the place of departure of the relative pronoun, as (ii) shows.

(ii) Ĵiske bareme betaiahe ki $\begin{Bmatrix} \text{vo} \\ \text{*—} \end{Bmatrix}$ bimarhe vo
whom about I told you that (he) sick is the
admi Ĵacukahe.
man has left.

The facts in (i) and (ii) require the ordering in (iii):

(iii) a. Relative pronoun copying;
b. Relative clause preposing (or, perhaps. relative head postposing);
c. Copy deletion.

Independently of ordering considerations, the facts in (i) and (ii) above are interesting, as they suggest that surface structure cannot be distorted without limit; thus, if the optional (iiib) applies, (iiic) is blocked, for the former creates decoding problems by causing the relative pronoun and the relative clause head to be non-adjacent, and the further perceptual difficulty which the deletion of the copy brings about can no longer be tolerated.

Hindi Relativization is subject to the CSC, as shown in (4.36). However, as shown by (4.37), this rule is not subject to the SSC.

(4.36) *me larka deikta hũ yiske bareme tumhe betaiahe
I boy see whom about I told you
ki John ketaakẽ nehi posand kerta aor Mary
that John books dislikes and Mary
posand karti he
likes
*'I see the boy about whom I told you that John likes books and Mary dislikes.'

(4.37) vo larka ǰise Mary čumti he tik he vo bimar he.
the boy who Mary kisses good is he sick is
*'The boy who that Mary kisses is a good thing is sick.'

4.2.4. The fourth case I wish to discuss concerns the possibility of deriving (4.38b) and (4.38a) (see Ross (1967, section 4.3)).

(4.38) a. The government prescribes the height of the lettering on the covers of the reports.
b. Reports which the government prescribes the height of the lettering on the covers of are invariably boring.

As I pointed out in section 3.2.4, sentences like (4.38b) constitute an apparent violation of the Subservient Satellite Constraint, as the moved element is not only a Satellite, but also an element of some other Satellite(s). By using (4.26), we may expect to explain the possibility of violating the Subservient Satellite Constraint in this case, since the interdiscontinuous material is indeed predictable by a simple recursive procedure roughly like (4.39).

(4.39) NP → NP NP_d [where the subscript on the second NP on the right of the arrow is an ad hoc abbreviation to which no theoretical significance should be attached indicating that that NP is a dependent of the preceding one].

This procedure is exceedingly simple, since it predicts a chain of NP's, each of which is a dependent of the NP immediately to its left.

We may also hope that (4.26) will explain the difference in acceptability between pairs like that in (4.40) (which have already been brought up in discussions of adverbials).

(4.40) a. The car which a reporter mentioned the scene of the crash of was a new Mustang.
b. *The car which a reporter found Mary at the scene of the crash of was a new Mustang.

Indeed, in both (4.40a) and (4.40b) we find a chain of NP's, in which every non-extreme NP is a Satellite of the NP to its left and a Nucleus of the NP to its right; but only in (4.40b) this whole chain of NP's is in turn a Satellite of a sentence, namely, *a reporter found Mary*. Thus, we may claim that even though both sentences exhibit a right-branching chain in which each pair of adjacent constituents exhibits the Nucleus-Satellite relation, it is only in (4.40a) that each Nucleus-Satellite pair is *of the same kind*; in (4.40b), the highest pair has the structure [S ADV], while the lower ones have the structure [NP NP]. Thus, it is only in (4.40a) that the chain of N & S's can be enumerated with *one* recursive procedure.

We have so far established that (4.38b) differs from other instances, in which the violation of the Subservient Satellite Constraint results in unacceptability, in that the maimed N & S is here recursively enumerable by one procedure; we have also established that the difference in acceptability between (4.40a) and (4.40b) correlates with the recursive predictability of the former, but not of the latter, with one procedure. However, in order to *explain* these facts, it is not sufficient to recognize that greater predictability requires less processing, and therefore less strain on immediate memory; some of the reasons are:

(a) Neither in (4.40a) nor in (4.40b) does the recursively predictable material constitute *the whole* of the interdiscontinuous material, and the processing of the latter requires more than one procedure in both cases. It seems that the crucial difference between the two sentences lies in whether *the chain of N & S's*, rather than *the whole of the interdiscontinuous material*, is recursively predictable with one procedure.

(b) As we shall see in 4.4, the unacceptability of structures which violate the Subservient Satellite Constraint cannot be (entirely) attributed to the structural complexity of the interdiscontinuous material and considerations of memory-load. To see this, we need only consider the difference in acceptability between (4.41a) and (4.41b).

(4.41) a. What did I tell you Bill likes?
b. *What did I tell you Bill likes and Mary
hates potatoes?

In (4.41a) and (4.41b), the interdiscontinuous material is identical, and therefore equally complex to process, but only the latter exhibits a violation of the Subservient Satellite Constraint.

Similarly, in (4.42), the only element occurring in the interdiscontinuous material which creates unacceptability is the head NP the claim.

(4.42) Who did you hear (*the claim) that John loves?

But the head of a complex NP need not occur between the discontinuous components. Thus, in Japanese, the head of a complex NP follows the modifying clause, though leftward migration of elements of that clause by Topicalization is out, as shown by (4.17d, e) and (4.17f, g).

(4.17) d.(*)Otto wa, watakusi wa kono boosi o kabutte
Otto I this hat wearing
ita to iu syutyoo o sinzita.
was that say claim believed.
*'Otto, I believed the claim was wearing this
hat.'

e. Watakusi wa Otto ga kono boosi o kabutte ita
I Otto this hat wearing was
to iu syutyoo o sinzita.
that say claim believed.
'I believed the claim that Otto was wearing
this hat.

f. Kore was Mary ga kabutte ita koto ga akiraka
This Mary wearing was thing obvious
no boosi da.
is hat is
'This is the hat which it is obvious that Mary
was wearing.'

g. *Mary wa, kore wa kabutte ita koto ga akirawa
Mary this wearing was thing obvious
na boosi da.
is hat is
*'Mary, this is the hat which it is obvious was
wearing.'

Notice also that in (1.36c), the adverbial reading of I suppose is out, due to the Subservient Satellite Constraint, but the Nucleus (John is sick) is not part of the interdiscontinuous material (there is, in fact, no discontinuity here).

(1.36) c. It's me who supposes that John is sick.

I conclude that the facts of predictability uncovered in relation to (4.38) and (4.40), even though probably relevant to the data at issue, cannot provide an explanation based solely on considerations of memory-load, and that something like the following suggests itself:

(4.43) A recursive procedure does not generate
significant new relations beyond those that
result from one of its applications.

What (4.43)--whose statement should undoubtedly be further constrained--purports to say is intuitively clear: (4.39) generates a second NP which bears the relation "nominal complement of" to the first NP, and no matter how many times the procedure applies, the lower nominal complements will still be perceived as *nominal complements only, not as parts of "higher" nominal complements*; in some sense, the application of the procedure (4.39) cannot generate new relations *just because it has applied more than once*. This is what we would expect if the structural hypothesis about chains like *the covers of the reports* and *the height of the lettering of the covers of the reports* is made essentially in the same way, i.e., by *one procedure* in both cases. The gist of the matter is that, given (4.43), (4.39) and (4.40a) are felt as instances of Satellite-chopping, but *not* of Satellite maiming. On the other hand, the moved constituent in (4.40a) is both a Satellite of the type 'nominal complement of NP' *and* part of a Satellite of the type 'optional adverbial'. Therefore, (4.40b), but not (4.40a), constitutes a violation of the Subservient Satellite Constraint, because the right-branching chain of N & S's is not recursively enumerable with a single procedure, as required by (4.43).

We can now reconsider the discussion in 4.2.1, in the light of (4.43). (4.22b) suggests that an object complement clause can be maimed, regardless of its 'depth'. In other words, a complement clause 'feels the same' regardless of whether it was generated by one application of (4.23) or more than one application, so that both (4.22b) and (4.44) are perceived as instances of maiming of *complement clauses* and not of instances of maiming of *different kinds of clauses*.

(4.44) It's John I believe Jill loves best.

It is tempting, on the basis of the above facts, to propose a universal principle like (4.45).

(4.45) If a discontinuity can involve a constituent of
type C generated by one application of a
recursive procedure PR, it can also involve
a constituent of type C generated by any
number of uninterrupted applications of PR.

I do not, at this stage, firmly propose (4.45) as a universal of language, since I do not quite know how to account for counterexamples like (4.46).

(4.46) a. John is upset [because Mary left him [because
he is crazy]].

(4.46) b. It's because Mary left him because he is crazy that John is upset.
c. It's because he is crazy that John is upset because Mary left him.

(4.46a) is a chain of right-branching N & S's generated by a recursive procedure (that the procedure is indeed recursive can be seen by continuing (4.46a) as ...because he fell on his head recently because somebody tripped him...). Because-clauses can be clefted, as shown by (4.46b), a 'cleft counterpart' of (4.46a). However, (4.46c) is not a cleft counterpart of (4.46a), as the clefted constituent can only be understood as modifying John is upset because Mary left him, but not Mary left him.

It appears that the lower because-clause in (4.46a) is felt as a part of a Satellite, not just as a Satellite. Obviously (4.46a) is, in some sense, different from the source of (4.22b) or from (4.38a), but stating the difference in a satisfactory way is no easy task. One differentiating factor could be the fact that the string in (4.46a), unlike the source of (4.22b) or (4.38a), is potentially ambiguous in many ways, since any of the because-clauses can be understood as modifying not just the clause immediately to its left, but any of the clauses to its left in the chain. Some of these possibilities are not realized when a clause other than the one immediately to the right of the main clause is clefted, since clauses modifying the same clause are logically coordinated, and clefting one of them is prohibited by the CSC. However, in (4.46c), at least two readings are theoretically possible: the clefted constituent logically modifies either John is upset because Mary left him or Mary left him. If it can be argued that the former reading is psychologically more complex than the latter (after all, the former is clearly more complex even in (4.46a), since it requires us to understand that John is upset for a very special reason, that is, that Mary left him because he is crazy, not just that she left him; this kind of complexity increases considerably as the chain is lengthened), then we are faced with a 'transderivational constraint situation' similar to the one discussed in 2.2.4.1 in connection with Perlmutter's a woman hit a girl who was pregnant; in other words, the resolution of a discontinuity is waived when the option of a less complex interpretation is available.

In all probability, these facts are only a partial explanation of the non-existence of a transformational relation between (4.46a) and (4.46b). In any event, I am hopeful that when a satisfactory explanation becomes available, it will be possible to uphold (4.45) as a universal of language, the non-derivability of (4.46c) from a structure incorporating (4.46a) being then attributable to well supported independent factors.

4.2.5. Before concluding section 4.2, which has been essentially concerned with the effect of predictability on structural complexity, it is perhaps worthwhile to illustrate the following rather obvious point: no complexity (in fact, no structural constraints of any kind) should arise when the frame in which a discontinuity is embedded is structurally irrelevant. Thus, consider (4.47).

(4.47) a. Somebody$_i$ met a man who loves someone's$_j$ wife.
b. Who met a man who loves someone's wife?
c. *Who met a man who loves whose wife?
d. Who met a man who loves Bill's wife?

We notice that (4.47b) but not (4.47c), can be derived from (4.47a). The reason is that in (4.47c) a feature-changing rule--wh-placement--violates the Subservient Satellite Constraint. However, (4.47c) is all right as an echo-question, when the interlocutor has said something like (4.47d) and the speaker of (4.47c) did not catch the word Bill; or as an incredulity-question, when the speaker finds it hard to believe that his interlocutor should ask that question about Bill's wife (possibly because he knows that she has been dead for many years, or that Bill is a bachelor, etc.). Observe, however, that echo- or incredulity-interpretations of (4.47c), the whole structure except for whose is structurally irrelevant, as it is only supplied as a framework for the hearer to be able to pinpoint the object of the illocutionary force (request to repeat a phrase, or expression of surprise at the presuppositions of that phrase respectively). As the structural frame is not intended for processing, we are confronted with the limiting case of structural non-complexity, namely, structural irrelevance, and consequently no constraints of the kind discussed in this thesis obtain.

A similar situation arises in the case of the so-called 'examination-type questions', that is, questions which are not genuine requests for information--since the speaker knows the answer, and his purpose in asking them is only to test his interlocutor's knowledge. Such questions can violate island constraints, as shown by (4.47e).

(4.47) e. Napoleon was a French general who invaded which country.

The grammaticality of (4.47e) on the examination-reading is explainable in essentially the same way as its grammaticality on the echo- or incredulity-reading. Indeed, everything in (4.47e) except for the underscored part is assumed to be common knowledge to both examiner and examinee; I suggest that the non-underscored part of (4.47e) is in fact part of the answer expected of the examinee (which the examiner provides, since it is known anyhow), and the illocutionary force of the utterance is something like 'complete the following answer by fitting the correct phrase in place of the one marked with a wh-feature'. Again, the non-underscored portion functions as a framework, rather than material meant to be processed.

4.3. This section is devoted to the Subservient Satellite Constraint. For reasons made clear below, the explanation will be based on conflict, rather than on interrupted behavior.

We have already seen in 4.2.4 that the unacceptability found in instances of violation of the Subservient Satellite Constraint cannot very well be attributed to the structural complexity of the interdiscontinuous material, since the material on which unacceptability could be blamed (that is, the Nucleus) is not necessarily part of the interdiscontinuous material. Moreover, it is implausible that the occurrence of a short and structurally simple Nucleus like *the fact* or *the claim* could place a load on memory sufficient to make all the difference between acceptability and unacceptability.

The first objection presented above is comparatively easy to take care of. Thus, rather than regard the Nucleus as the unpermissible 'intervening material', we shall attribute that status to the N & S node, by extending the notion *intervene* to apply not only to the linear, but also to the vertical, dimension (thus, we say that a node A which dominates B but not C *intervenes* between B and C). We can see that regardless of the relative linear ordering of the Nucleus and the Satellite, violations of the Subservient Satellite Constraint will always create a discontinuous Satellite with an intervening N & S node, as the element external to the Satellite is *by definition* outside the dominance of the N & S node.

Concerning the second objection, I conclude that the complexity of interrupted behavior can indeed not be regarded as essentially responsible for the Subservient Satellite Constraint, and that a more plausible account can be given in terms of perceptual conflict. Specifically, I propose that the following principle is (at least in part) responsible for the Subservient Satellite Constraint:

(4.48) A stimulus may not be perceived as simultaneously having *maximally* conflicting values on the same classificatory scale.

In the Introduction, I pointed out that a generalization like (4.49) would constitute a near-truism, and therefore hardly present any interest; as can be seen, (4.48) is considerably more specific than (4.49) (this was also true of the conflict principle (2.23)).

(4.49) Conflicting percepts are complex.

Before showing how the Subservient Satellite Constraint follows from (4.48), I will briefly argue that such a principle is independently needed for linguistic phenomena. Consider Langacker's restriction on the occurrence of pronouns:

(4.50) A pronoun may not bear *all* relevant primacy relations to its antecedent.

The relevant scale here is one of *relative emancipation*. Thus, when some notion or referent occurs more than once, one occurrence is usually regarded as 'primary' while the others are regarded as 'mere repetitions'. The primary occurrence can be identified in at least two ways: positionally and morphologically. The positional relations involve the two available dimensions: *precede*, for the horizontal dimension, and *command*, for the vertical one (to convince ourselves that the relation *command* expresses a genuine psychological reality, we may think of asymmetrically commanding elements as analogous to commanding officers in a military unit; the grammatically irrelevant relation 'higher up in the tree than' is analogous to the militarily irrelevant one 'superior officer of'). In most situations, both positional relations are relevant; in symmetric coordinations, only precede is, as the terms are hierarchical equals of one another. In both situations, when all the positional evidence points in one direction, the morphological evidence cannot point in the opposite direction, which is precisely what (4.50) says formally. Observe, however, that partial conflict is, if not ideal, at least tolerable. Thus, (4.51a), in which the morphological evidence is contradicted by the horizontal, but not vertical, evidence, is an acceptable sentence, although perhaps less felicitous than (4.51b), in which no conflict obtains.

(4.51) a. The man who said that he loves her announced yesterday that he would marry Julie soon.
b. The man who said that he loves Julie announced yesterday that he would marry her soon.

In all probability, (4.50) is a grammatization of (4.48), and thus provides internal evidence in its favor. I do not know of specific corroborating experimental evidence (the complexity of conflicting percepts has, of course, already been studied extensively). In arguing in favor of (4.52), in fact, a weaker subcase of (4.48), Bever (1970) mentions supporting evidence from visual perception.

(4.52) A stimulus may not be perceived as simultaneously having two positions on the same classificatory dimension.

Bever argues that (4.52) is responsible for the complexity of multiple center-embedding, of triple negation, of chains of recursively right- or left-branching N & S's, etc. I do not believe that (4.52) is solely responsible for the unacceptability of the constructions Bever considers, as those constructions differ sharply in degree of acceptability (for some arguments that independent factors are involved see 2.1.3). Nevertheless, (4.52) does seem to be at least partly responsible for the complexity of triple negation, where the medial negative is both 'negator' and 'negatee', or for the (rather faint, in my opinion) complexity of the picture of the owner of the house, where the owner is a dependent of the picture but the head of the house. (4.52) is undoubtedly a much weaker restriction than (4.48), since it involves an apparent, rather than a real, conflict; indeed, the medial element in the cases Bever considers is emancipated with respect to some element and downgraded with respect to another element, the only problem being that the two (or more) pairs of elements are judged with respect to the same relation; in contrast, in cases which violated (4.50), an element (the pronoun) is both maximally emancipated and maximally downgraded with respect to the same element (the antecedent).[8] Consequently, it is to be expected that violations of (4.48) will be much more severe than violations of (4.52), and that grammatizations of the former will in general occur.

---

[8] The definition I proposed for symmetric coordinate terms as bidirectional N & S's (when there are two terms) does not predict

Having established the psychological reality of (4.48) in language perception, we turn to the relation between (4.48) and the Subservient Satellite Constraint.

Hierarchical relations among constituents also exhibit a scale of *relative emancipation*, as we have seen coreferents do. In endocentric constructions in general, the head bears a primacy relation to its dependents, which we shall call the relation of *centrality*. In N & S's, the Nucleus is not only central, but also self-sufficient, since the Satellite, unlike other types of dependents, is not a partner in the N & S, but an optional addition; we shall say that the Nucleus bears to the Satellite the relation of *sufficiency*. These two relations transfer quite naturally to the N & S node in relation to the Satellite, the latter being *non-central* and *non-essential* with respect to the former. Let non-centrality and non-essentialness be *status relations*. Then in terms of such status relations, a Satellite is maximally downgraded with respect to its N & S node.

Let us now consider what the psychological import of a discontinuous Satellite is, when the N & S node intervenes in the sense defined above. In a situation of that nature, the structural arrangement is such that the well-formedness of the Satellite is not decidable without taking the external element into account; in other words, the structural integrity of the Satellite depends on something else than the N & S node or the material it dominates. Thus, on the scale of structural emancipation, the relevant status cues suggest that the Satellite is wholly dependent on the N & S node, while the structural-integrity cues suggest the opposite; hence, a conflict which results in complexity, as (4.48) predicts.

Despite superficial differences, the Subservient Satellite Constraint is quite similar to Langacker's constraint on pronominalization. In both cases, total conflict is unacceptable and partial conflict is not. Thus, analogous to the acceptability of (4.51a), we find the endocentric constituent VP, which may consist of a verb and

---

that symmetric coordinations are complex, as (4.48) is not applicable. Indeed, I did not propose that the terms are *simultaneously* perceived as Nuclei (emancipated) and (Satellites (downgraded), but rather that we can look at symmetric coordinations in as many ways as there are terms. Put differently, in the symmetric structure,

$$\left[_{X}\ X_1 \quad \left[_{X} \begin{Bmatrix}\text{and}\\ \text{or}\end{Bmatrix} \quad X_2\right]\right]$$

X_1 is Nucleus and X_2 is Satellite, but we know that for each such structure there is a synonymous one differing structurally from this one only in that X_2 appears in the position occupied by X_1 and vice versa.

a complement clause (as well as other elements not relevant to our discussion); in such a case, the dependent, i.e., the complement clause, is non-central with respect to the VP node, but it is not non-essential; consequently, the maiming of the complement clause is permissible.

Notice that the conflict between status and integrity facts which arises due to violations of the Subservient Satellite Constraint cannot arise when the whole Satellite has been chopped. Indeed, as the Satellite does not exist under the domination of the N & S node, it is patent that its structural integrity cannot be decided on the basis of the material dominated by the N & S node, and the opposite assumption does not arise.

Before concluding this section, it should be pointed out again that violations of the Subservient Satellite Constraint are not always observable from surface structure; this fact has always been recognized, and it led Ross to his claims about rule ordering (Ross (1967)), and later scholars to the reformulation of the various island constraints as global ones. At least four cases can arise in which the surface structure does not contain all the necessary information; they are illustrated in (4.53)-(4.56) below.

(4.53) *It's Mary who I believe what you told.

(4.54) *This the girl I saw a boy who likes.

(4.55) *It's Mary who a boy just left who likes.

(4.56) (?)I am aware of the fact that you like someone,
but I don't know who.

In (4.53), the intervening N & S node is not obviously present in surface structure, since the Nucleus has been deleted; in (4.54), the external element which forms one end of the discontinuity (the relative pronoun) has been deleted; in (4.55), it is not obvious we are dealing with an N & S, as the Nucleus and the Satellite are non-adjacent; finally, in (4.56), the whole N & S is absent, the only suggestion of a discontinuity at some level being the external element who. In the first three cases, the degree of unacceptability is essentially the same as in the corresponding overtly clear cases (in fact, (4.55) is probably worse, in view of the difficulty involved in integrating a double discontinuity); in the last case, unacceptability is considerably reduced for most speakers, and non-existent for some speakers.

Regarding (4.53), I argued at length in 4.1.3.1 that the reconstruction of heads on the basis of the double function of words like wh-forms is a highly efficient strategy, and therefore the hearer will have, in all probability, reconstructed the N & S node at the point at which integration is attempted.

Concerning (4.54), it is not really true that the external element is invisible; instead, the external element is now the head of the relative clause (the girl), rather than the wh-pronoun which could have appeared next to it. The situation is similar to the one found with Japanese relative clauses, where the relative pronoun is always invisible (in 4.4, I shall argue that the distinction between Japanese and English is a trivial one).

With respect to (4.56), the improvement in acceptability due to the disappearance of most of the N & S (as well as of additional material, in certain cases) deserves a separate study, which I have not undertaken. The first temptation is to say that speakers do not necessarily interpret (4.56) by reconstructing the whole of the deleted material, but in some ad hoc way, by appending something like he is at the end of (4.56). Another possibility would be to assume that even if the deleted material is reconstructed, it need not be reconstructed with a deletion site in place of the moved element, but rather with a pronominal copy (we recall that copying rules are not subject to the Subservient Satellite Constraint). Chomsky (1970b) suggests that 'understood material' as in (4.56) should be interpreted by 'compositional semantics', rather than by postulating a deletion. This proposal, like the guesses I have made above, is insufficiently elaborated on to be amenable to discussion.

Any serious proposal along the above lines should explain why forbidden operations which 'take place inside deleted strings' sometimes cause unacceptabilities and sometimes do not. Thus, unacceptabilities due to violations of Cross-Over are still felt when the evidence that there has been a permutation of coreferents is no longer present in the string, as shown by (4.57).

(4.57) *Mary was hit by Bill$_i$, but he$_i$ wasn't [hit by himself$_i$].

It might be suggested that there is no way in which compositional semantics could reconstruct the bracketed part of (4.57) without producing a coreferent of he$_i$, while the missing string in (4.56) can be reconstructed without actually producing a maimed island. But until more pertinent data are considered and until these vague proposals are made precise, there is little point in further speculation.

4.4. In Chapter One, I mentioned Ross' observation that certain rules do obey his island-constraints, while others do not. His statement, which I called the Dichotomous Behavior Principle (DBP), is reproduced below.

(1.52) Chopping rules, feature-changing rules, and unidirectional rules of pronominalization obey the constraints; bidirectional rules of pronominalization and copying rules do not.

I have pointed out in various places in this thesis that the DBP is inadequate for at least the following two reasons:

(4.58) a. Island constraints cannot be restricted to transformations.
b. Ross' distinction between unidirectional and bidirectional pronominalization is both false and misleading.

In relation to (4.58), Ross' distinction is false, because he attributes the unacceptability of (4.59) to the fact that VP-Deletion (a *bidirectional* rule) goes into a complex NP (as we have seen, deletion is in fact irrelevant).

(4.59) *John $\left\{\begin{array}{l}\text{is taller than}\\ \text{as tall as}\end{array}\right\}$ I know a boy who is.

Furthermore, Ross' distinction is misleading because it misses the point that deletion rules are subject to island-constraints if and only if *there is an obligatory identity condition between the 'deletor' and the 'deletee'*. The Japanese rule of Relativization is a case in point, and some of the rules of 'unidirectional pronominalization' Ross mentions constitute illustrations of the same principle. The rules illustrated in (4.60) are particularly interesting because the coreferentiality condition between the underscored elements in (4.60a) and (4.60b) is invited, in the sense of Geis and Zwicky (1971), rather than real.

(4.60) a. *This rock* is too heavy for me to try to help Bob to pick (*it*) up.
b. The *socks* are ready for you to put (*them*) on.

The inference-inviting factor in constructions like those in (4.60) is the fact that the *for-to* clauses are 'resultative', in the sense that the events or state of affairs expressed in them must be a consequence of the state of affairs expressed in the main clause. Consider, for example, (4.61).

(4.61) a. This rock is too heavy for John to try to lift the table.
b. The socks are ready for you to put on the bathrobe.

Upon being presented with sentences like (4.61a) or (4.61b), many informants reject them as semantically incoherent. However, when some of the unexpressed assumptions are pointed out to them, they realize that the sentences in question are semantically well-formed. Thus, if it is assumed that John is bearing the rock on his shoulders while trying to lift the table, (4.61) makes perfect sense; the state of affairs expressed in the *for-to* clause now follows from that assumption in conjunction with the proposition expressed by the main clause. Similarly, if we imagine a culture in which one is forbidden to put on a bathrobe unless a pair of socks hangs from the chandelier, the *for-to* clause of (4.61b) follows from that piece of pragmatic information in conjunction with the main clause, and the whole sentence becomes entirely coherent.

It should be noted that, while the *for-to* clauses in (4.61a) or (4.61b) do not contain a coreferent of *the rock* or *the sock* respectively, coreferents of these phrases must show up in the pragmatic presuppositions of the two sentences, which we may formulate as (4.61c) and (4.61d) respectively.

(4.61) c. John is bearing the rock on his shoulders.
d. The socks must hang from the chandelier before one is entitled to put on a bathrobe.

I do not see how the pertinent assumptions of (4.61a) and (4.61b) could be coherently formulated unless they contain an occurrence of the rock or the socks respectively. Consequently, a coreferentiality condition must be satisfied at some level for sentences like (4.60a) and (4.60b) or (4.61a) ~~and~~ (4.61b) to make sense. In the absence of assumptions like (4.61c) and (4.61d), a hearer assumes that the necessary connection between the main for-to clause in sentences like (4.60a) and (4.60b) will be made explicit in the actual string, and expects a coreferent of the subject of the main clause to show up in the for-to clause. Put differently, the resultative character of the for-to clause strongly invites the inference that it will contain a coreferent of the subject of the main clause. This invited inference is sufficient to cause the rule which may delete the coreferent inside the for-to clause to be subject to essentially[9] the same constraints as rules which delete a coreferent under strictly linguistic conditions of obligatory coreferentiality, such as Japanese Relativization. Ross (1967) illustrates the point with respect to the first part of the CSC, and the LBC and the SSC, in the case of the deletion rules manifested in (4.60); as none of these constraints are subcases of my Subservient Satellite Constraint, I provide pertinent illustrations in (4.62) and (4.63). A parallel paradigm is constructed for Japanese Relativization in (4.64).

(4.62) a. The rock is too heavy for me to believe that a boy willing to pick $\left\{\begin{matrix}\text{it}\\ {}^{*}\_\_\_\end{matrix}\right\}$ up exists.

b. This rock is too heavy for me to both lift Mary and pick $\left\{\begin{matrix}\text{it}\\ {}^{*}\_\_\_\end{matrix}\right\}$ up.

c. This rock is too heavy for me to believe that you broke your back in lifting $\left\{\begin{matrix}\text{it}\\ {}^{*}\_\_\_\end{matrix}\right\}$ up.

---

[9] A. Zwicky informs me that the deleted versions of the a-sentences in (4.62) and (4.63) are out for him, but that the c-sentences are marginal, while the b-ones are OK. I don't know what to make of this fact, except assume that the obligatory coreferentiality condition may be less strong for some speakers when it is only invited; however, I do not know how to account for the gradations in acceptability within the above paradigms which Zwicky reports.

(4.63) a. The bathrobe is ready for you to find a girl
willing to put $\left\{\begin{matrix}\text{it}\\ *\_\_\end{matrix}\right\}$ on.

b. The bathrobe is ready for you to simultaneously
take $\left\{\begin{matrix}\text{it}\\ *\_\_\end{matrix}\right\}$ off and put your socks on.

c. The bathrobe is ready for you to get decent
by putting $\left\{\begin{matrix}\text{it}\\ *\_\_\end{matrix}\right\}$ on.

(4.64) a. *Suki na onno no ko ga sotta teeburu wa Tookyoo
Likes girl left table tokio
ni aru.
in is
*'The table which the girl who likes left is in
Tokyo.'

b. *John wa Mary ga suki de, Jim ga suki na onna
John Mary likes Jim likes
no ko ma kiree da.
girl pretty
*'The girl who John likes Mary and Jim likes
is pretty.'

c. *John ga suki ta kara Tookyoo e itta onna no ko
John likes because Tokio to went girl
wa okotte iru.
mad is
*'The girl who John went to Tokyo because he
loves is mad.'

It can be seen that (4.62)-(4.64) and (4.59), which Ross lumped together as unidirectional pronominalization phenomena, are essentially different; thus, unacceptability arises in the latter case independently of deletion, while in the former, it arises only when deletion, not merely pronominalization, has applied. Also, Ross' only reason for regarding the deletion phenomena in (4.60)-(4.64) as unidirectional is that the complement clause of heavy or ready apparently cannot be preposed; but, by the same token, there is no reason for claiming that the rule which allows a non-null pro-form in the above examples is bidirectional, and if it is not, then non-null and null pronominalization appear to behave differently. Notice that Ross cannot claim that that rule of non-null pronominalization is the same as pronominalization in general, for he chose to regard VP-Deletion in comparatives and equitatives as unidirectional and therefore different from VP-Deletion in general.[10]

---

[10] Actually VP-Deletion in comparative or equitative clauses differs from VP-Deletion in other environments in being obligatory, as the contrast between (i) and (ii) shows.

(i) John is clever, but Bill isn't $\left\{\begin{matrix}\text{clever}\\ \_\_\_\_\end{matrix}\right\}$.

(ii) John is taller than Bill is $\left\{\begin{matrix}\text{*tall}\\ \_\_\_\_\end{matrix}\right\}$.

I do not think, however, that this fact should be construed as an indication that the rule which optionally deletes *clever* and the one which obligatorily deletes *tall* in (i) and (ii) respectively are distinct rules.

Given the inadequacies of the DBP, I propose to reformulate it as the observationally more adequate (4.65):

(4.65) The DBP': The Subservient Satellite Constraint is applicable to chopping, deletion under obligatory coreferentiality, feature-changing, and semantic neutralization; it is not applicable to non-null pronominalization or to deletion without obligatory coreferentiality.[11]

The explanation I propose for (4.65) is (4.66).

(4.66) The situations in which the Subservient Satellite Constraint applies are perceived as involving a structural discontinuity; the remainder are not.

The point of the claim made by (4.66) can be seen most clearly by comparing chopping and copying phenomena, which the DBP' assigns to different subclasses, as shown in (4.67a) and (4.67b) respectively.

(4.67) a. *This is the girl *who* you met *a boy who likes*.
b. This is the girl *who* you met *a boy who likes her* [dialectally restricted].

In both sentences, the N & S and the pertinent external element have been underscored. The essential difference is that the Satellite is ill-formed in (4.67a) unless integrated with the external element, while this is not so in (4.67b). In this connection I propose the following hypothesis:

(4.68) The relation which obtains between the external element and the Satellite in sentences like (4.67a) and (4.67b) are processed by *essentially different strategies*.

---

[11] I mention deletion under obligatory *coreferentiality* because I do not know of any deep structure conditions mentioning obligatory *identity of sense*.

Thus, in the former case, the discontinuity is resolved by structural integrative strategies; in the latter, it is handled by strategies which recognize semantic relations between elements. The recognition of identity of reference in (4.67b) is, I suggest, not essentially different from the recognition of a relation of antonymy or 'opposition' between the underscored constituents in (4.69a) and (4.69b) respectively.

(4.69) a. John is clever, but his cousin is stupid.
b. John doesn't dislike cake, he loathes it.

I hope that it is intuitively clear why I do not regard the sentences in (4.69) as containing structurally discontinuous elements, although it is not easy to make this notion formally precise. Also, I do not know of any experiments which support or cast doubt on (4.68), and it is not easy to see how one would go about setting up pertinent experiments. I hope, however, that my proposal is a plausible one, and I will try to show in the remainder of this section that it provides a natural and internally consistent explanation for the DBP'.

Notice that it would be possible, in principle, for sentences like (4.67a) to be interpreted by reconstructing a null anaphor in the position of the object of like, which would be later interpreted as coreferential with who by semantic strategies. If that were the case, we would expect there to be no distinction in acceptability between (4.67a) and (4.67b), but the fact that there is such a distinction suggests that a null anaphor is not reconstructed. The reason is, I suggest, that the recognition of the existence of a necessary coreferent of some element 'alerts' the integrative strategies, so that integration will have to take place when the chance arises; on the other hand, nothing alerts the strategies which reconstruct missing elements, and what takes place is integration, rather than reconstruction followed by recognition of identity.

The position I have adopted predicts that deletion under obligatory identity will always behave identically, for the purposes of the Subservient Satellite Constraint and other island types, with chopping rules. As far as I know, this prediction is borne out by the facts (see (4.62)-(4.64)). A corollary of this position is that whenever the deletion of a pro-form is exempt from island constraints, there is no obligatory condition that a candidate for deletee status exist; I will show below that this corollary is borne out by the available facts. Finally, this position predicts that Neubauer's claim (which I argue against in Appendix Four) to the effect that chopping rules behave differently from copying rules followed by the deletion of the original cannot be correct in principle, for the effect of copying is to introduce a necessary coreferent of the original.

Notice also that for a theory which regards constraints like the Subservient Satellite Constraint as 'purely grammatical' ones, there is no way to capture the similarity between chopping and deletion under obligatory identity. Indeed, if chopping proceeds in one step, the Subservient Satellite Constraint will be a transformational constraint for chopping and a global one for deletion; if chopping proceeds in two steps, the Subservient Satellite Constraint will be a global rule in both cases, but it will have to mention the output of copying in one case and underlying representation in the other.

Before leaving the subject of chopping and deletion under obligatory identity, I think it worthwhile to dispose of an apparent counterexample to my account.

(4.70) *This is the apple *which* I know *a boy who ate*.

In (4.70), the N & S is not incomplete, since we can understand it as *a boy who ate something*. Unlike (4.67), where I claimed that nothing alerts the reconstructive strategies, the latter *are* alerted in (4.70) by the lexical properties of *eat* (which allows the deletion of its object if that object is a *generic*). As reconstructive strategies are generally based on 'local cues', i.e., found in the same syntactic perceptual unit as the missing element, and since perceptual units must undergo some processing before they can be effectively integrated, one would expect reconstructive strategies to apply before integrative ones in general. Given such an ordering of strategies, the unacceptability of (4.70) seems to pose a problem. The problem is, however, a spurious one: the missing elements in (4.70) can only be reconstructed as a generic, and *generics have no antecedents*. Therefore, if a generic *is* reconstructed, the sentence fails because the semantic strategies cannot establish a connection between it and *which*; if a generic is *not* reconstructed, the sentence fails for the same reasons as (4.67a).

In contrast to (4.70), consider the following paradigm (taken from Ross (1969b)) which shows that the reconstruction of an *anaphoric*, rather than *generic*, pro-form does not lead to unacceptability.

(4.71) a. *I know that he must be proud of it, but I don't know how. [the sentence is unacceptable only on the reading 'I don't know how proud of it he must be', not on the reading 'I don't know in which way he must be proud of it'].

b. Johnny stole someone's wallet, but I forget whose.

The difference in acceptability between (4.71a) and (4.71b) is due to the fact that there exists a rule which deletes the heads of genitives under identity of sense, as in *I like this book, but I don't like Bill's (book)*, while there is no rule which deletes the heads of degree modifiers, as can be seen in (4.72).

(4.72) Jill is pretty, but I don't know how $\left\{\begin{matrix}\text{pretty}\\ {}^{*}\text{\_\_\_\_}\end{matrix}\right\}$ she is.

Consequently, the inapplicability of a head-constructive strategy to (4.71) makes it impossible to recover it as (4.73a), while the applicability of such a strategy to (4.71b) makes possible its construction as (4.73b).

(4.73) a. I know that he must be proud of it, but I don't know how proud (of it he must be).
b. Johnny stole someone's wallet, but I forget whose $\left\{\begin{matrix}\text{wallet}\\ \underline{\hspace{2em}}\end{matrix}\right\}$ (he stole).

Notice that neither (4.73a) nor (4.73b) violate the LBC, while (4.71a) does, which explains its unacceptability.[12] The contrast in acceptability

---

[12] G. Drachman has pointed out to me that (i) is marginally acceptable in his speech.

(i) (?)I know that he must be proud of it, but I don't know how much.

This suggests to me that (i) must be derived from something like (ii), because how much is not in general movable from a left branch, as shown in (iii).

(ii) I know that he must be proud of it, but I don't know [$_s$ he must be proud of it to wh-some extent].

(iii) a. John has a lot of money.
b. How much money does John have?
c. *How much does John have money?

An apparent counterexample is furnished by (iva), since its source seems to include (ivb), rather than the ungrammatical (ivc).

(iv) a. You don't know how very much John is in love with Mary.
b. John is very much in love with Mary.
c. *John is in love with Mary very much.

However, the ungrammaticality of (iiic) suggest that (ivc) is, after all, part of the source of (ivc), and (ivc) must become (ivb) by an obligatory preposing rule.

I should like to point out that, while the LBC does indeed account for (4.71a), there are certain problems with it which suggest that some reanalysis would be in order.

Thus, if how may be dominated by the category NP, it is hard to extend this categorization to prepositions and articles; still, such a move is suggested by the impossibility of deriving (vb) from (va) or (vib) from (via).

(v) a. John is ín love with Mary [not óut of love with her].
b. *It's ín that John is love with Mary.

(vi) a. I saw thé man, not á man.
b. *What I saw man was thé, not á.

Notice that the LBC may be insufficient even if prepositions and articles are categorization as NP's, for it is highly improbable that languages with postpositions or post-nominal articles allow the movement of such constituents.

This would seem to require a RBC, instead of the LBC, in such languages. Notice, however, the Rumanian has both pre- and post-nominal articles (the indefinite and definite ones respectively), but if we impose both the LBC and the RBC in Rumanian, we wrongly predict that sentences like (vii) or (viii) are ungrammatical.

(vii) Cît e Petre de deştept!
How is Peter of clever
'How clever Peter is!'

(viii) Diferenţa e enormă între un cerc şi un pătrat.
'The difference is enormous between a circle and a square.

I should also point out that Ross' explanation of the freezing of the heads of Complex NP's by the LBC strikes me as incorrect, for heads do not move in Japanese either (see (4.104) and (4.105)), where they appear on the right-branch of higher NP's, or in Hindi, where they may occur on either the right- or the left-branch of higher NP's.

---

between (4.71b) and (4.70) is thus due to the fact that the empty slot in the former is interpreted as a null anaphor, equivalent to one, whose antecedent is wallet, while the empty slot in the latter can only be interpreted as a generic, and therefore non-anaphoric, or as an empty slot, in which case the integrative strategies are alerted.

We have so far discussed chopping, deletion under obligatory coreferentiality and non-null pronominalization (illustrated by copying). It remains to discuss the inapplicability of the Subservient Satellite Constraint to deletion without obligatory identity, and its applicability to feature-changing and deep structure neutralization.

With respect to deletion without obligatory identity, the inapplicability of the Subservient Satellite Constraint follows quite straightforwardly from the preceding discussion, as the lack of an obligatory identity condition fails to alert the integrative strategies. In (4.74a)-(4.74d), I illustrate the non-applicability of the CNPC to the rules of Genitive-Head Deletion, Super-Equi-NP-Deletion, VP-Deletion and Sluicing respectively.

(4.74) a. I know that you like Bill's book, but you must face the fact that I am going to promote Greg's.
b. Max was worried that the fact that it was dangerous to shave himself with his rusty old blade bothered Shirley.

(4.74) c. Just try to square a circle, if you believe the guy who told you that you can!

d. Somebody put ice cream in Jim's bed and He didn't believe the guys who told him that they didn't know who.[13], [14]

In none of the sentences of (4.74a)-(4.74d) is there a condition *requiring* identity between the deletion-controller and the deletion site, as (4.75a)-(4.75d) respectively show.

(4.75) a. I know that you like Bill's book, but you must face the fact that I am going to promote Greg's play.

b. Max was worried that the fact that it was dangerous for Bill to shave himself with his rusty old blade bothered Shirley.

c. Just try to square a circle, if you believe the guy who told you that you can use a pencil adequately.

d. Somebody put ice cream in Jim's bed and he didn't believe the guys who told him that they didn't know who had tried to get even with him.

Concerning feature-changing phenomena, the applicability of the Subservient Satellite Constraint is entirely expected. Indeed, Satellites containing forms like *any*, *ever*, *wh*-pronouns, etc., are not well-formed unless commanded by the appropriate neutralizer; unacceptability arises when the latter is outside the dominance of the N & S node, since the structural well-formedness of the Satellite can only be determined by recognizing an unpermissible discontinuity.

---

[13] Examples (4.74a)-(4.74d) are from Neubauer (1970).

[14] The rules illustrated in (4.74) supply additional evidence that the two parts of the CSC are different phenomena, being subcases of the Frozen Nucleus Constraint and the Subservient Satellite Constraint respectively. Indeed, while these rules are not subject to the latter constraint, they are subject to the former. I illustrate the point with Genitive-Head Deletion below; parallel paradigms for the other rules can easily be constructed by the reader.

(i) I know that you like Bill's books, but I am going to promote Greg's $\left\{\begin{array}{c}\text{books}\\ \_\_\_\end{array}\right\}$ and veto everything Bill has ever written.

(ii) I know that you like Bill's books, but I prefer John's $\left\{\begin{array}{l}\left\{\begin{array}{l}\text{books and plays}\\ {*}\_\_\_\end{array}\right\}\\ \left\{\begin{array}{l}\text{plays and books}\\ \qquad\quad\;\; {*}\_\_\_\end{array}\right\}\end{array}\right\}$

Finally, the semantic neutralization of the 'comparee' by the 'comparator' in comparative and equitative constructions is subject to Subservient Satellite Constraint for the same reasons as feature-changing; the well-formedness of the clause containing the comparee can only be established by inspecting the comparator. This is shown by the contrast in acceptability between (4.76a) and (4.76b), which have the same adverbial clause, and where acceptability depends on the compatability of the comparator and the comparee ((4.76c) shows that (4.76b) contains a possible comparator).

(4.76) a. Joanne is more intelligent than Jill is efficient.
b. *Joanne loves her husband more than Jill is efficient.
c. Joanne loves her husband more than Jill hates her brother.

In (4.76a) and (4.76c), the comparee occurs inside a Satellite (the comparative clause), but the N & S node does not intervene (in the sense discussed above), since the comparator is precisely the Nucleus of this N & S. On the other hand, the sentences in (4.77) are all bad, because an N & S node does intervene between the comparator and the comparee, and the latter is part of the corresponding Satellite.

(4.77) a. *Mary is $\left\{\begin{matrix}\text{more}\\ \text{as}\end{matrix}\right\}$ intelligent $\left\{\begin{matrix}\text{than}\\ \text{as}\end{matrix}\right\}$ I know a boy who is productive.
b. Mary is $\left\{\begin{matrix}\text{more}\\ \text{as}\end{matrix}\right\}$ intelligent $\left\{\begin{matrix}\text{than}\\ \text{as}\end{matrix}\right\}$ you heard (*the claim that) she is productive.
c. *Mary is $\left\{\begin{matrix}\text{more}\\ \text{as}\end{matrix}\right\}$ intelligent $\left\{\begin{matrix}\text{than}\\ \text{as}\end{matrix}\right\}$ she is productive and an assistant professor.
d. *Mary is $\left\{\begin{matrix}\text{more}\\ \text{as}\end{matrix}\right\}$ intelligent $\left\{\begin{matrix}\text{than}\\ \text{as}\end{matrix}\right\}$ Bill is in love with her because she is productive.

Semantic neutralization as in (4.76)-(4.77) is subject to the Subservient Satellite Constraint in Japanese as well, as illustrated for the coordinate subcase in (4.78).

(4.78) *John wa taijuu ga 500 ponto te, tsuyoi yori Bill
John weight 500 pounds (and) strong than Bill
wa motto tsuyoi.
more strong
*'Bill is stronger than John weighs 500 pounds and is strong.'[15]

---

[15] It is interesting that while the semantic neutralization found in comparatives is subject to the Subservient Satellite Constraint in Japanese, feature neutralization is not, at least in the cases that I am acquainted with. Thus, Ross (1967) pointed out that Japanese Reflexivization is free from his constraints, and I illustrate the same point below with respect to the Japanese counterpart of *Wh*-Placement (i.e., the rule which converts an indefinite NP to *dare* when that NP is the 'target' of some question verb).

(i) Watakusi wa dare no sukina onna no ko ga Tookyoo e
I whom likes girl Tokyo to
itta no ka to kiita.
went asked.
*'I asked the girl who likes whom went to Tokyo.'

(ii) John wa Mary ga suki de, Bill wa dare ga suki desu ka.
John Mary like Bill who like
*'John likes Mary and Bill likes who?'

(iii) Watakusi wa John wa dare ga sukita kara Tokyoo e
I John who likes because Tokyo to
itta no ka to kiita.
went asked.
*'I asked John went to Tokyo because he loves who.'

It should be pointed out that most speakers of English have the feeling that, while violations of the Subservient Satellite Constraint by feature neutralization are unacceptable, they are not nearly as bad as the sentences in (4.77); for example, (iv) is considerably more tolerable than (4.77c).

(iv) *I wonder who likes Moscow and which other city?

This suggests that violations of the Subservient Satellite Constraint by the comparator-comparee relation are, in some sense, much stronger than similar violations by the relation holding between a neutralizer and a neutralized feature. In fact, it is not altogether unexpected that it should be so, for the neutralization of a comparee by a comparator is semantic, and therefore of greater impact than feature neutralization, which is essentially formal (notice that feature neutralization, as in (iv), while based on a semantic *relation* between the question verb and some *wh*-word, nevertheless involves no semantic *neutralization* of the latter by the former). However, it should be made clear that the sentences in (4.77), though bad enough to deserve a double asterisk, are by no means *semantically* ill-formed. For example, (4.78d) means that Bill is in love with Mary because she is productive and that she is productive to a lesser extent than (or to the same extent as) the one to which Mary is intelligent; thus, (4.78d) makes perfect sense semantically, and its unacceptability must be attributed to some non-semantic principle, such as the Subservient Satellite Constraint.

To summarize, I have argued in this section that the partition determined by the DBP' on syntactic phenomena depends on whether a structural discontinuity is or is not perceived. I proposed that the interpretation of Satellites which contain a pro-form whose antecedent is outside the N & S requires the recognition of the antecedent-anaphor relation, but this does not prevent the Satellite from being perceived as whole; similarly, deletion sites can be interpreted as null anaphors, provided that there is no indication that some external element necessarily belongs in that slot. With respect to the phenomena which are subject to the Subservient Satellite Constraint, I argued that they do involve a structural discontinuity; for chopping and certain deletions, because an external element must be integrated with the Satellite, and for feature and semantic neutralization, because the well-formedness of the Satellite is partly determined from outside the N & S.

4.5. In this section, we consider the problems raised by the Frozen Nucleus Constraint with respect to chopping (as I pointed out in Chapter Three, the facts are insufficiently clear for other types of phenomena to warrant an attempt at explanation at this stage; I suppose, however, that deletion under obligatory coreferentiality can be explained in essentially the same way as chopping).

Two kinds of situations can result from chopping the Nucleus and leaving a Satellite behind:

(4.79) a. The clause out of which the Nucleus was lifted appears well-formed, or
b. The clause out of which the Nucleus was lifted appears ill-formed.

I will try to show that both (4.79a) and (4.79b) are perceptually complex situations, involving erroneous closure and conflict respectively (as far as I can see, interrupted behavior does not seem to be crucially involved).

(4.79) is illustrated in (4.80) and (4.81). Contrast this paradigm with the one exhibited in (4.82) and (4.83).

(4.80) a. I believe (that) Mary's claim that Bill left is false.
b. *Whose claim do I believe (that) [$_s$ that Bill left is false]?

(4.81) a. I think that John, Bill and Mary have left.
b. *Who do I think that [$_s$ John and Mary have left]?

(4.82) a. The difference between X and Y is clear.
b. [$_s$ The difference is clear] between X and Y.

(4.83) a. You think that the roof of that house was blown off by a bomb.
b. Of which house do you think that [$_s$ the roof was blown off by a bomb]?

In both (4.80)-(4.81) and (4.82)-(4.83), the clause out of which an element has chopped seems well-formed, but the total result is bad in the former, where a Nucleus was moved, while it is acceptable in the later, where a Satellite was moved. I suggest that the unacceptability of the b-sentences in (4.80)-(4.81) is due to a principle of erroneous closure, rather similar to (2.8). Specifically, I claim that the unacceptability of (4.80b) and (4.81b) is due to the fact that the bracketed constituents are erroneously apprehended as complete sentences, and that they cannot be coherently integrated with the remainder of the string *without undergoing significant reanalysis*.

Thus, in (4.80), the string *that Bill left* is a possible subject of *is false*, and is interpreted as such; the correct integration of *whose claim* with *that Bill left* would require that the analysis of the latter as a subject of the bracketed clause be discarded, and that it be reinterpreted as a modifier of another NP which is the real subject of that clause (in fact, the Nucleus of the subject NP, but Nuclei are by definition representative of their N & S's).

In contrast, even though the bracketed clauses of (4.82b) and (4.83b) are possible clauses, their integration with the displaced constituents *does not require that they be significantly reanalyzed*. For example, *the roof* in (4.83b) is essentially still the subject of the bracketed clause, even after the optional appendage *of which house* has been attached to it. Thus, the difference in acceptability between Nucleus-chopping and Satellite-chopping in (4.80)-(4.83) is due to the fact that the former, but not the latter, requires a significant reanalysis of the form in which some substring has been initially construed.

It should be pointed out that all the cases of erroneous closure so far considered involved the *significant reanalysis* (in the formally undefined, but intuitively clear sense of the above account) of the pertinent string; in fact, I suspect that complexity arises only when erroneous closure results in a percept which must be significantly reanalyzed at some later stage of processing. It would therefore be worthwhile to attempt to provide a rigorous definition of the notion 'significant reanalysis', without which the prediction I have just made cannot be formally stated.

We now consider case (b) of (4.79), which is illustrated in (4.84)-(4.86).

(4.84) a. You think those members of the committee are dead.
b. *Which members do you think [$_S$ of the committee are dead?

(4.85) a. I told you John and Mary have left.
b. *Who did I tell you [$_S$ and Mary have left]?

(4.86) a. That John was fired because he was sick is appalling.
b. *What [$_S$ because he was sick is appalling] is that John was fired.

(4.84)-(4.86) are different from (4.80)-(4.81) in that the bracketed clauses do not appear to be well-formed; we may expect this to function as a clear indication of a slot, thus making integration easy. However, (4.84)-(4.86) involve a dffficulty of a different sort; unlike (4.82)-(4.83), the remainder of the N & S's (i.e., of the committee, and Mary, and because he was sick respectively) are the wrong kind (in the sense of Appendix Two) of category for their being arguments of any sort of the predicates of the bracketed clauses, and the initial analysis of those clauses is rendered difficult by the perceptual conflict between the categorial kind of some constituent and the role it is called upon to play.

Like (4.79a), (4.79b) does not require new perceptual principles for its explanation. The unacceptability of (4.84b), (4.85b) and (4.86b) is predicted by (4.48): see section 4.3); as there are no additional categorial or functional cues that could be brought to bear with respect to the pertinent bracketed strings, we are faced with a situation of maximal conflict between the role that those strings must play and the categories to which they belong.

The fact that Nucleus-chopping results in a perceptually complex situation does not, of course, mean that no counterexamples can exist in principle. Consider thus (4.87b)[16] which exhibits a violation of the Frozen Nucleus Constraint.

(4.87) a. I saw someone interesting at the party.
b. Who did I see interesting at the party?

Notice that we cannot claim that someone interesting was moved integrally by Question-Movement, the reduced relative clause somehow extraposing later, for the latter need not appear at the end of a clause, as shown in (4.88b). The badness of (4.89b) is an additional argument against extraposition of a reduced relative, for a full relative can appear in that position, as shown by the acceptability of (4.89a).

(4.88) a. You showed someone interesting to my girl friend from Rome.
b. Who did you show interesting to my girl friend from Rome?

(4.89) a. Who did you meet yesterday who was interesting?
b. *Who did you meet yesterday interesting?

To begin with, notice that the class of adjectives which can be substituted for interesting (4.87b) or (4.88b) is fairly restricted; thus, it includes devastating and frightening, but not pretty.

---

[16] I am grateful to D. T. Langendoen for calling this case to my attention.

repugnant, surprising, appalling, or astonishing (the adjectives which behave like interesting do not seem to form a natural class, but I have not looked at the problem in detail; moreover, there seems to be serious idiolectal variation in this respect).

Secondly, it can be shown that the structures found in (4.87b) or (4.88b) are complex, since acceptability decreases rapidly when the chopped elements belong in more deeply embedded clauses. Thus, compare the marginal status of (4.90a) and the unacceptability of (4.81a) with the acceptability of (4.90b) and (4.91b), although the interdiscontinuous material is identical in each member of each set.

(4.90) a. ??Who did I tell you Mary met interesting?
b. Who did I tell you Mary met who was interesting?

(4.91) a. *Who did I tell you Bill claims Mary met interesting?
b. Who did I tell you Bill claims Mary met who was interesting?

However, why can the Nucleus in constructions like someone interesting be chopped at all? The answer is, I believe, that it must, rather than can, be chopped, in order to avoid another unacceptable construction. Indeed, there exists a constraint (whatever its explanation) which prohibits the sequence who interesting, as shown in (4.92a). Thus, sentences like (4.87) seem to result from a trade-off between constraints, such that the conflict is resolved by the (partial) suppression of one of them.[17] Certain adjectival modifiers are better than others after an interrogative pronoun, and my suggestion of a trade-off between constraints is supported by the paradigm in (4.92)-(4.94), where Nucleus-chopping seems to increase in acceptability as N & S-chopping decreases in acceptability.

(4.92) a. *Who interesting did you show (to) the audience?
b. Who did you show interesting to the audience?

(4.93) a. ?Who $\left\{\begin{matrix}\text{with}\\\text{wearing}\end{matrix}\right\}$ a red scarf did you show (to) the audience?
b. ??Who did you show $\left\{\begin{matrix}\text{with}\\\text{wearing}\end{matrix}\right\}$ a red scarf to the audience?

---

[17] Of course, there is no reason in principle why both constraints could not have held (the meaning of (4.87b) can be conveyed by using a full, rather than reduced, relative clause), and I am prepared to believe that there are dialects in which both Constraints hold, with the result that both (4.87b) and (4.92a) are out.

(4.94) a. Who who was interesting did you show (to) the audience?
b. *Who did you show who was interesting to the audience?

It is important to distinguish genuine violations of the Frozen Nucleus Constraint from spurious ones. Thus, there are a number of rules which move Satellites to the end of the containing clauses, and when the N & S originates in clause-final position, its chopping followed by rightward migration of the Satellite looks as if the Nucleus alone had been chopped. Such situations will not cause perceptual problems, because the Nucleus will be synthesized with an optional slot for a Satellite (as discussed above), and the Satellite will not be in a position which conflicts with its categorial kind (since clause final position is ambiguous between some argument of the predicate and 'extraposed' constituent, provided that there is a rule of rightward migration relating the Nucleus and the Satellite). One such case of spurious Nucleus-chopping was recognized in Ross (1967), where it was pointed out that sentences like (4.95c) arise from (4.95a) through the intermediate stage (4.95b), rather than directly from (4.95a).

(4.95) a. I saw a man who was wearing a tie.
b. Who who was wearing a tie did I see?
c. Who did I see who was wearing a tie?

Other instances of N & S chopping followed by Satellite migration have been misconstrued. For example, in Stockwell et al., (1968), it is claimed that the heads of postnominal genitival constructions like the underscored constituent in (4.96a) can be clefted, yielding (4.96b). I believe that clefting yields the intermediate stage (4.96c). from which (4.96b) is derived by the independently motivated rule of Extraposition-of-PP. My reasons for thinking so are that (4.97b) is not derivable from (4.97a), although we would expect (4.97b) to be grammatical if Nucleus-clefting were possible.

(4.96) a. I found a necktie of John's.
b. It's a necktie I found of John's.
c. It's a necktie of John's that I found.

(4.97) a. I showed a necktie of John's to Mary.
b. *It's a necktie I showed of John's to Mary.
c.?*It's a necktie I showed to Mary of John's.

(4.97c) is, unfortunately, bad for independent reasons (extraposition movements cannot cross a prepositional phrase in general). The badness of (4.97b) is, however, a sufficient reason for doubting that (4.96b) arises by Nucleus-chopping.

D. Perlmutter (oral communication) has claimed that NP's in construction with ves 'all' can be chopped in Slovenian. His argument was that (4.98) must be a case of genuine head-chopping, for ves exhibits the ordinary accusative ending expected if concord takes place between it and riz prior to chopping; on the other hand, if the head of ves had been copied and deleted through an intermediate stage of pronominalization to ga (it), the quantifier would have agreed with ga, exhibiting the animate accusative ending, as in (4.99).

(4.98) Riž sem želel projesti ves.
The rice I wanted to eat all

(4.99) Želel sem ga pojesti vsega.
I wanted it to eat all

Notice, however, that the above facts do not support Perlmutter's claim in case there is a rule of Quantifier-Floating in Slovenian, similar to the one exhibited in (4.100) for English.

(4.100) a. All the boys have left.
b. The boys all have left.
c. The boys have all left.
d.?*The boys have left all [this sentence has become ungrammatical quite recenlty; it was perfectly acceptable in 19th century English].

As I am not acquainted with Slovenian, I cannot provide an analysis of (4.98), but I will show that similar facts obtain in Rumanian, and that they can be explained there by an independently motivated rule of Quantifier-Floating similar to the one found in English.

Thus, consider (4.101) and (4.102).

(4.101) a. Am văzut Parisul tot.
I saw Paris all
'I have seen all of Paris'.

b. Parisul, l-am văzut tot.
Paris, it I saw all
'Paris, I have seen all of it.'

(4.102) a. Am văzut Romînia toată
I saw Rumania all
'I have seen all of Rumania.'

b. Romînia, am văzut-o toată.
Rumania I saw it all
'Rumania, I have seen all of it.'

As we can see, the quantifier tot 'all' exhibits the same kind of agreement before and after topicalization. However, this quantifier can also float quite freely, as shown by (4.103). Consequently, sentences like (4.101b) and (4.102b) can be derived from their a-counterparts in two stages, by topicalizing the entire N & S and then allowing the quantifier to float.

(4.103) a. Toţi băieţii au plecat.
'All the boys have left'
b. Băieţii toţi au plecat.
'The boys all have left'
c. Băieţii au plecat toţi.
'The boys have left all'

The conclusion which may be drawn from the above discussion is that once the real and/or apparent counterexamples have been accounted for in a principled way, the Frozen Nucleus Constraint emerges as a genuine principle of language, which can receive a straightforward explanation in terms of general and independently motivated principles of performance.

In contrast, consider Ross' account of the freezing of Nuclei. He only discusses this problem in relation to complex NP's, and notes, in section 2.3, that the chopping of the head of a relative clause or of a complement clause 'will be shown to be excluded by either of two independent conditions: the Complex NP Constraint... or the Pied-Piping Convention'. As the CNPC is formulated in his 4.1, it cannot freeze the heads of complex NP's, for it only refers to elements of the modifying clause; therefore, it appears Ross decided to leave it to the Pied-Piping Convention (and, specifically, to the Left Branch Condition on the latter) to block the migration of the heads of Complex NP's.

I believe that it is wrong to account for this phenomenon with the LBC (see also note 12), because of the implied claim that the heads of complex NP's cannot move *only when they precede the modifying clause*. This conclusion strikes me as unnatural in the extreme; moreover, it is probably false. Thus, the head noun of complex NP's always follows the modifying clause in Japanese, but they cannot be chopped either. (4.104b) and (4.105b) show that Japanese Topicalization, which moves constituents to sentence initial position, cannot move the head of a relative clause or a complement clause respectively.

(4.104) a. Kodomo go tabete iru sakana wa ookii.
child eating is fish big
'The fish which the child is eating is big.'
b. *Sakana wa kodomo ga tabete iru ookii.
fish child eating is big

(4.105) a. Otto ga kono boosi o kabutte ita to iu
Otto this hat wearing was that
syutyoo o watakusi wa sinzita.
claim I believed
'I believed the claim that Otto was wearing this hat.'
b. *Syutyoo wa, Otto ga kono boosi o kabutte ita
the claim, Otto this hat wearing was
to iu o watakusi wa sinzita.
that I believed.
*'The claim, I believed that Otto was wearing this hat.'

I believe we can conclude that the freezing of the heads of complex NP's is much more adequately accounted for by the Frozen Nucleus Constraint than by the LBC, since the former, but not the latter can account (a) for the freezing of complex NP heads which _follow_ the modifying clause, and (b) for the freezing of Nuclei in constructions _other than complex NP's_.

CHAPTER V

SUMMARY AND FURTHER OUTLOOKS

The purpose of this dissertation has been twofold:

(5.1) a. To provide an improved syntactic analysis of the phenomena known as 'island constraints', and

b. to offer a natural explanation of those and other phenomena in behavioral terms.

In regard to (5.1a), I believe the major contribution of this dissertation to be the notion Nucleus and Satellite, formally definable in terms of the notions dominate and categorial kind; there is overwhelming evidence that the latter two are independently needed in an adequate description of natural languages, and a large part of my effort has been directed towards showing that the notion of N & S must also be part of a descriptively adequate grammatical theory. The notion N & S has made it possible to recognize that the notion of island constraint was a cover term for two largely independent and weaker constraints, the Frozen Nucleus Constraint and the Subservient Satellite Constraint; moreover, the excessive power of the various island constraints forces the recognition of a large number of counterexamples, an embarrassment which my proposal, on the whole, avoids.

With respect to its second purpose, this dissertation has attempted to show that the phenomena discussed in Ross (1967), Postal (1971), and elsewhere, can be accounted for in terms of behavioral principles involving erroneous closure, interrupted behavior, and perceptual conflict. Closure principles underlie my account of some instances of the Frozen Nucleus Constraints, as well as of some instances of the Right Roof Constraint; interruption principles underlie my accounts of some instances of the Cross-Over phenomena involving unbounded rules, of the Right Roof Constraint, of the optionality of Pied-Piping, of the possibility of moving elements out of arbitrarily deeply embedded clauses, and of the problems created by the migration of elements out of pre-verbal clauses in SVO languages; conflict principles underlie my accounts of the Cross-Over phenomena involving bounded rules, of some instances of the Frozen Nucleus Constraint, of the Subservient Satellite Constraint, and of the Langacker-Ross constraint on backwards pronominalization.

It has become a commonplace nowadays to end a paper by pointing out that it has raised many more questions than it has answered; given our extremely inadequate knowledge of the mental reality underlying syntactic phenomena, such an observation is almost a truism.

As far as this dissertation is concerned, there are countless pertinent issues it has not even attempted to tackle, such as:

(5.2)
a. What is a possible behavioral principle?
b. How do behavioral constraints interact with semantic ones in determining acceptability?
c. Are there linguistic phemomena which can be explained in terms of speech production, but not of speech perception?
d. Is it possible to predict grammatization, and if so, what are the nature, the weights, and the principles of interaction, of the pertinent variables?
e. Are there any 'purely syntactic' universal constraints?

Questions like the above are of great theoretical interest, but I do not think the appropriate answers will be forthcoming in the near future.

I believe that problems of more immediate interest concern the work which remains to be done in connection with the issues which this dissertation has attempted to tackle. First, as my research has centered largely on English (and possibly not even on all the pertinent areas of English), my proposals are seriously underdetermined by the data, and additional research on a reasonable sample of the languages and dialects of the world is an absolute must. Second, it is certainly desirable to submit as much as possible of my analysis to experimental testing, for, even if my proposals are plausible ones, they are certainly not the only logically possible ones. If this dissertation succeeds in stimulating such work in the future, I will consider that it has accomplished its purpose, even if every single claim it has made should eventually turn out to be wrong.

APPENDIX ONE

I do not wish to claim that (2.56) is the deepest representation of (2.54a), but only a remote representation of it, because (2.56) is not semantically interpretable. Various ways of dealing with the semantic interpretation of (2.56) suggest themselves, the most satisfactory of which seems to be (a) below.

(a) At the deepest level, the boxed NP is empty, exactly like the Predicate, and the semantic interpretation of this structure is carried out by the same rule which interprets the underlying representation of the synonymous (2.54) (the empty nodes as well as the copula would be disregarded). This rule of interpretation is the deep structure counterpart of the surface interpretive rule proposed in Chomsky (1969): while in the latter account, heavy stress is assigned optionally and the presupposition is obligatorily inferred in the presence of stress, in my account, presupposition may or may not exist, and heavy stress is obligatorily assigned in the presence of presupposition.

The presupposition of a sentence must be a proper subpart of the meaning of that sentence. Thus, (A1.1a) may presuppose not only (A1.1e), but also (A1.1b, c, or d).

(A1.1) a. John saw Mary.
b. John saw a girl.
c. John saw a person.
d. John saw someone.
e. John saw something.

Corresponding to (A1.1b)-(A1.1e), the pseudo-cleft output will be (A1.2b)-(A1.2e) respectively.

(A1.2) b. The girl who John saw was Mary.
c. The person who John saw was Mary.
d. The one who John saw was Mary.
e. What John saw was Mary.

Similarly (A1.3a) may presuppose (A1.3b), in addition to (A1.3c).

(A1.3) a. I noticed John's grisly sense of humor.
b. I noticed something about John.
c. I noticed something.

Corresponding to (A1.3b) and (A1.3c), we have the pseudo-clefts (A1.4b) and (A1.4c).

(A1.4) b. What I noticed about John was his grisly sense of humor.
c. What I noticed was John's grisly sense of humor.

Sentences like (A1.4b) have been claimed by Green (1971) to have no non-pseudo-cleft counterpart and thus to create serious problems for any analysis of pseudo-cleft constructions which includes a non-pseudo-cleft counterpart. However, given universal principles of semantic implication, the non-pseudo-cleft counterpart of (A1.4b) is (A1.4a), not the ungrammatical **I noticed John's grisly sense of humor about him* which Green regards as the source of the difficulty.

In slightly more complicated situations, the presupposition is inferrable from the non-pseudo-cleft counterpart in conjunction with a number of pragmatic assumptions (for some discussion, see Lakoff (1971)).

Consider (A1.5a) and the pragmatic assumptions (A1.5b)-(A1.5e).

(A1.5) a. A man like Hitler was allowed to seize power.
b. Hitler was a vicious tyrant.
c. The world has a duty to prevent vicious tyrants from seizing power.
d. The world did not prevent Hitler from seizing power.
e. When one does not do one's duty, something is wrong with the one who thus fails.

From (A1.5b)-(A1.5e), we can infer (A1.6).

(A1.5) Something is wrong with the world.

Taking (A1.5a) as the non-pseudo-cleft counterpart and (A1.6) as the presupposition, we can construct the pseudo-cleft in (A1.7).

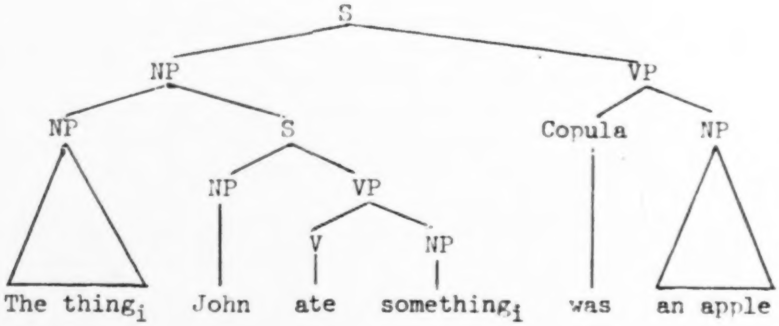
(A1.7) What's wrong with the world is that a man like Hitler was allowed to seize power.

Returning now to the simpler (2.54a), the remote representation (2.56) is derived by copying *something* (i.e., the pro-form corresponding to the focus) into the boxed NP, and definitizing it later because of the 'previous mention' (that definite pronouns and determiners can have their antecedents in presuppositions is shown by sentences like *the boy saw Mary*, which is meaningless, unless a previous mention of the referent of *the boy* is presupposed).

If the rule which inserts a copy of the pro-form into the boxed NP seems insufficiently motivated, consider the alternatives:

(b) We may regard (2.56) as a pre-semantic structure, in which case the semantic interpretation would apply to the output of Focus-Placement (see below), that is, (A1.8).

(A1.8)



Presupposition could also be determined at this stage (on the basis of the general principle that restrictive relatives with definite determiners are presupposed), and so could coreferentiality. However, apart from the fact that singling out the output of Focus-Placement for semantic interpretation would be no less arbitrary than the copying rule I have proposed, this solution could not work because Focus-Placement has to follow the rule which determines the distribution of *any* (see the discussion in connection with examples (2.58) and (2.59), which in turn has to follow Passive, as (A1.9) shows:

(A1.9) a. *Anybody didn't kill John.
b. John wasn't killed by anybody.
c. John didn't kill anybody.
d. *Anybody wasn't killed by John.

But if Focus-Placement has to follow Passive, it would be wrong to single out its output for semantic interpretation, for, in case the Passive has applied, passive and active structures would have to be interpreted by separate rules, thus missing an obvious generalization.

(c) Another possibility would be to regard (i) as essentially the deepest representation of (2.1), as proposed in Postal (1971). Such an analysis would create no problem for semantic interpretation, but would entail the loss of all the syntactic generalizations about pseudo-cleft constructions which will be discussed below, namely the correct statement of selectional restrictions, the distribution of neutralized elements, the impossibility of removing elements from pseudo-cleft phrases, and the fact that the acceptability of pseudo-cleft constructions is determined by the principles which limit rightward movements i.e., (2.35c). Moreover, to derive sentences like (2.54b), Postal is forced to regard something like (2.54a) as a more remote representation and to posit a rule of Contrast-Movement. That rule would have to apply obligatorily to (A1.10a) to derive (A1.10b). The ill-formedness of (A1.10a) appears as an accidental fact in this analysis.

(A1.10) a. *What I saw Bill the table was under.
b. I saw Bill under the table.

On the other hand, if (A1.10b) is regarded as more basic than

(A1.10a), the latter is excluded by the general principle that transformations do not move 'function words' (or parts of words, as pointed out in Chomsky (1969)).

A still more serious problem is raised by (A1.11a) which one could say in reply to (A1.11b), but for which, given Postal's analysis, it is hard to imagine a source.

(A1.11) a. No, Jóhn saw the hóuse.
b. Jill saw the table.

Given the untenability of (b) or (c) above on syntactic grounds, I have assumed an analysis along the lines of (a) in 2.1.3.1, because the awkwardness of having empty nodes in underlying structure is not an intolerable difficulty, and because (a) makes it possible to handle the syntactic problems involved in pseudo-clefting, which is a primary concern of that section.

APPENDIX TWO

Optionality is in fact not important for its own sake, but rather for showing that tne N & S-node is necessarily of the same 'categorial kind' as in the Nucleus-node. The notion of 'kind' that I have in mind is ill-defined in linguistic theory at present, since the only notation available for labeling categories does not differentiate between gerundials, that-clauses and for-to-clauses, for example, all being assigned the label S. Nevertheless, there is a great deal of syntactic evidence that finer categorial distinctions are needed, since transformations are sometimes sensitive to degrees of sentencehood, nominality, or adjectivalness (for details, see Ross (1972)). The failure to recognize such degrees has been responsible, I believe, for a great deal of unnecessary argumentation in the literature. For example, Emmonds (1969) argued that gerundials are NP's, not S's, because they do not extrapose; on the other hand, gerundials behave like sentences with respect to other syntactic phenomena, such as pronominalization constraints. Thus, unless we regard gerundials as clauses it is hard to account for the ungrammaticality of (A2.1).

(A2.1) *He$_i$ disapproves of your flattering John$_i$.

The evidence, however, no longer appears contradictory if we posit a continuous scale ranging from nominality to sentencehood, such as in (A2.2).

(A2.2) SENTENCEHOOD
a. that-clauses
b. For-to-clauses
c. gerundial clauses
d. gerundial nominalizations
e. deverbal nouns
NOMINALITY

A categorial scale like (A2.2) imposes an implicational constraint, in the sense that every rule or constraint which mentions the category S or NP must partition the scale into two continuous sets, one of which may be null. That is, the correctness of (A2.2) would be challenged by a constraint which would block (A2.2b) and (A2.2d), but not (A2.2a, c, and e). I do not know of such counterexamples, and I give a few examples below which support (A2.2). (A2.3)-(A2.5) show (A2.2) is confirmed by the Sentential Subject Constraint, the NP-over-S constraint, and Chomsky's Specified Subject Condition (see Appendix Three) respectively.

(A2.3) a. *It's John who that you killed is clear.
b. *It's John who for you to meet may be dangerous.
c. *It's John who meeting may prove dangerous.
d. It's John who the shooting of has become imperative.
e. It's John that a picture of is in my pocket.

(A2.4) a. *Is that John left obvious?
b. *Is for John to leave desirable?
c. Is John's leaving the country desirable?
d. Is John's shooting of so many rabbits legal?
e. Is John's picture of Mary available to you?

(A2.5) a. It's John who I was told that you met at the party.
b. It's John who I intend for you to meet.
c. *It's John who I am planning on Mary's meeting.
d. *It's John who I am planning on a detective's shooting of.
e. *It's John that I saw Bill's picture of.

Two more phenomena which support (A2.2) are Extraposition and Extraposition-from-NP, as shown in (A2.6) and (A2.7) respectively.

(A2.6) a. It's desirable that John should leave.
b. It was expected for John to leave.
c. *It's regrettable John's leaving.
d. *It's regrettable John's shooting of the * rabbits.
e. *It was unexpected John's proof of the theorem.

(A2.7) a. The claim is preposterous that John is a traitor.
b. *The hope is unrealistic for John to get well.
c. *The prospects are dim of John's getting well.
d. *The possibility is frightening of the shooting of so many rabbits.
e. *The idea never occurred to me of the proof of the theorem.

Concerning (A2.7e), it should be pointed out that there are sentences like (A2.8), which apparently violate the hierarchy in (A2.2).

(A2.8) The choice is hard between A and B.

Thus, unlike (A2.7e), (A2.8) is acceptable, even though the constituent extraposed from NP is a PP (which is presumably at the nominal end of the scale). However, it seems reasonable to assume that (A2.8) is not derived by Extraposition-from-NP, but by another rule (usually called Extraposition-of-PP). Although I have not investigated the domain of applicability of the latter rule, it seems clear that it is more restricted than that of the former, and I suspect that detailed investigation would turn out additional formal differences between the two rules.

Another reason for believing that the two rules should not be collapsed is that there are acceptable cases of 'Extraposition-of-NP'; that is, exactly as there are grammatical sentences apparently derived like (A2.7e), there also are grammatical sentences like (A2.9) which are apparently derived like (A2.6).

(A2.9) It's awful the way he behaves in front of women.

That (A2.9) is not derived by ordinary Extraposition is argued in detail in Elliot (1971, Ch. 2), where it is shown that there is a separate rule of Extraposition-of-NP restricted to 'exclamatory' predicates.

If Extraposition and Extraposition-from-NP are not involved in the derivation of (A2.9) and (A2.8) respectively, the claim that these two rules recognize (A2.2) is consistent with the facts.

The notion of categorial kind as illustrated by (A2.2) is clearly not the only possible or existing one. Although I do not have anything like a theory of categorial kinds at present, I feel it is rather obvious that a prepositional phrase and the noun phrase it necessarily contains are not NP's of the same kind. For example, the doubly-underscored NP in (A2.10a) cannot be substituted for the singly-underscored one, as (A2.10b) suggests.

(A2.10) a. I did it for John.
b. *I did it John.

Similarly, although coordinations are represented as in (3.99), so that the boxed X and X_2 belong to the same category, it is quite clear that they are not of the same kind, for the former cannot appear in most contexts in which the latter can, as illustrated in (A2.11).

(A2.11) a. Bill left.
b. *Or Bill left.

The ungrammaticality of (A1.10b) and (A2.11b) is a consequence of the obligatory character of prepositions in PP's and of or in second disjuncts, as opposed to the optional character of non-restrictive relative clauses in NP's. Therefore, even if the notion of categorial kind could be made explicit in linguistic theory, replacing (3.1) by another definition (3.1') which would differ from (3.1) in lacking provision (iii) and in having a strengthened provision (i) such that X and Y would be required to belong to the same kind of category would be no more than proposing a notational variant, so long as optionality is the only reason for claiming that X and Y are of the same kind.

Notice also that (3.1) is trying to capture the intuition that it is a significant fact about N & S's that their Nucleus is of the same kind as the entire construction. Thus (A2.12) contains two NP's of the same kind, which are moreover substitutable in the context given; however, the fact that the two NP's are of the same kind is not a necessary or significant one, and (3.1) correctly prevents the internal NP from being regarded as the Nucleus of the larger one.

(A2.12) A man who was talking to a woman had a heart attack.

APPENDIX THREE

The fact that some cases of Extraposition-from-NP are awkward does not indicate that (3.2iii) has questionable status; rather, the awkward cases can be explained by independent constraints. We have already seen two such constraints (in Chapter Two): (a) the acceptability of extraposed clauses decreases in proportion to the complexity of the material which intervenes between them and their original head, and (b) all other things being equal, a reading of a sentence that involves Extraposition-from-NP tends to be discarded in favor of another reading which involves no extraposition or extraposition over less complex material. These constraints were shown in Chapter Two to be quite general and in no way restricted to the rule in question. I wish to show now that (c) Extraposition-from-NP yields increasingly unacceptable results in proportion to the degree of specificity of the head NP. Thus, consider (A3.1).

(A3.1) a. A man just left who was wearing a hat.
b. ?The man just left who was wearing a hat.
c.??That man just left who was wearing a hat.
d. *John's brother just left who was wearing a hat.
e.**John's brother's son's daughter just left who was wearing a hat.

The constraint which is operative in (A3.1) is clearly more general than meets the eye; indeed, as was originally pointed out by Warshawsky (1966) and recently elaborated on in Chomsky (1971), the complements of NP's are increasingly impervious to operations involving the complement and an element external to the N & S in proportion to the specificity of the determiner of the Nucleus. Thus, the reflexivization or chopping of a nominal complement exhibit the same gradual deterioration as that seen in (A3.1), as shown by (A3.2) and (A3.3) respectively (additional phenomena which behave similarly are discussed in Chomsky (1971)).

(A3.2) a. I gave John a picture of himself.
b. ?I gave John that picture of himself.
c. *I gave John Mary's picture of himself.
d.**I gave John Mary's sister's daughter's picture of himself.

(A3.3) a. What did I give you a picture of?
b. ?What did I give you that picture of?
c. *What did I give you Bill's picture of?
d.**What did I give you Bill's father's brother's picture of?

Chomsky proposes to handle these cases by his Specified Subject Condition, which I reproduce below:

(A3.4) No rule can involve X, Y, in the structure
...X...[α ...Z...-WYV...]...
where Z is the specified subject of WYV in α.

This condition in effect states that a discontinuous relation cannot cross a specified subject, but (A3.1) shows that the crossing is irrelevant. Moreover, (A3.1) which is intuitively felt to be a subcase of whatever constraint is operative in (A3.2) and (A3.3) cannot be handled by (A3.4). I suggest that a more observationally adequate statement would be (A3.5).

(A3.5) In N & S's belonging to the category NP, (3.2iii) becomes increasingly inapplicable in proportion to the specificity of the Nucleus.

As shown by (A3.2), the increasing inapplicability of (3.2iii) concerns not just chopping, but feature-changing as well (similar facts hold for the distribution of any, as shown in Ross (1967)). In fact, (A3.5) affects the same rules as those affected by (3.2iv); as shown in (A3.6), (A3.5) affects deletion under obligatory coreferentiality, but not non-null pronominalization. Copying rules are not affected either (as illustrated in (A3.7) and (A3.8) with Left Dislocation and WH-Q-Copying respectively, but further research is necessary, as I have not been able to check the acceptability of (A3.8) with enough informants who question by copying).

(A3.6) Mary is too ugly for John to want to own
$\left\{\begin{array}{l}\text{a picture of}\\ \text{this picture of??}\\ \text{Ed's picture of*}\\ \text{Bill's son's picture of**}\end{array}\right\}$

It should be noticed that the entire set in (A3.6) becomes acceptable if her is added at the end of each sentence.

(A3.7) My father, I think I mentioned $\left\{\begin{array}{l}\text{a}\\ \text{this}\\ \text{Bill's}\\ \text{John's father's}\end{array}\right\}$ picture of him.

(A3.8) Who did I show you $\left\{\begin{array}{l}\text{a}\\ \text{that}\\ \text{Bill's}\\ \text{Mary's mother's}\end{array}\right\}$ picture of him?

The above facts suggest that (A3.5) should be replaced by something like (A3.9).

(A3.9) The phenomena which cannot affect elements of a Satellite, as predicted by (3.2iv), are prevented from affecting the entire Satellite, (a) if the Nucleus is an NP, and (b) in proportion to the specificity of the Nucleus.

If (A3.9) is correct, then it must be taken into account when explanations for (3.2iv) are considered, for it is suggested rather strongly that the fact that the _same_ restrictions hold for elements of Satellites and for whole Satellites under conditions of specificity is more than a mere coincidence.

Finally, notice that (A3.9) explains the often-noted fact that many speakers prefer to extrapose relative clauses rather than sentences in apposition to NP's like _the fact_. Indeed, the determiner of the Nucleus in the latter case can often not be less specific than _the_, in view of the ill-formedness of _*a fact that John is crazy is disturbing_ (see section 3.2.3), and Nuclei having _the_ as determiner tend to inhibit extraposition, regardless of the kind of clause to be extraposed.

APPENDIX FOUR

In this Appendix, I consider a possible explanation of (3.110) in terms of (3.2). My suggestion is that (3.110) is in fact unnecessary, since whatever purpose it purports to serve is essentially served by (3.2iv). Thus, (3.2iv) only precludes the migration of an element of a Satellite from under the dominance of the N & S node, but does not preclude movements within the domination of the N & S node; for example, relative pronouns can be attracted to their antecedent. A number of non-chopping phenomena which are subject to (3.2iv) create no unacceptabilities when they involve an element internal to the Nucleus and another internal to the Satellite; this is illustrated with ever-neutralization, deletion under obligatory coreferentiality, and comparative neutralization in (A4.1)-(A4.3) respectively.

(A4.1) I know no one who ever shot two lions with one shot.

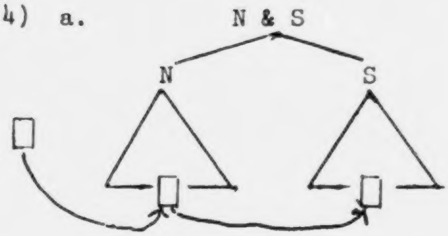
(A4.2) The man (who) you mentioned has left.

(A4.3) John loves Mary more than he hates Jill.

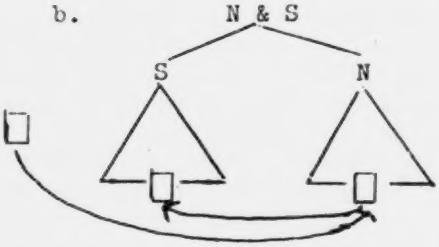
To explain (3.110) we need only to assume that when the Nucleus and the Satellite of an N & S contain an identical element in 'similar' positions, the element in the Satellite is first moved into the Nucleus where it is 'gobbled' by the identical element; the latter is subsequently moved out of the N & S. Thus, I am positing sequential, rather than simultaneous, movements in across-the-board processes. Gobbling rules of the type suggested above have been proposed in various places in the literature; the syntactic justification of my proposal is that it avoids the embarrassment of regarding across-the-board movements as exceptions to (3.2iv). This account makes sense especially in behavioral terms; indeed, if the reconstruction of discontinuities created by across-the-board movements proceeds as in (A4.5), a perceptual conflict will arise (see section 4.3), but if it proceeds as in (A4.4a) or (A4.4b), no difficulty will arise.

A minor apparent theoretical difficulty concerns the seemingly arbitrary choice between (A4.4a) or (A4.4b).in symmetric coordinations. The real-time procedure is probably (A4.4a), since the first term carries no mark of subordination, and there consequently seems to be no reason for the reconstructive procedure to hold the first term unresolved until the second is encountered. Similarly, in n-term coordinations, reconstruction very probably proceeds from left to right.

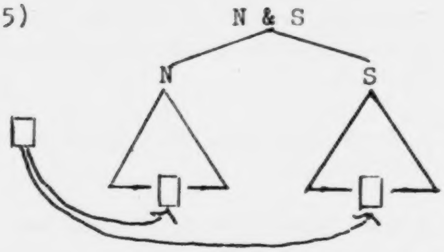
(A4.4) a.



b.



(A4.5)



An alternative way of dealing with across-the-board maiming while preserving the CSC is implicitly proposed in Neubauer (1970). Neubauer's analysis is interesting to consider because it makes the claim that copying rules are exempt from the constraints which affect chopping rules, even when copying is followed by the deletion of the original, so that the end product is indistinguishable from chopping.

Neubauer's argument hinges on the fact that Conjunction-Reduction can remove elements from coordinations, unlike Question-Movement, Relativization, Pseudo-Clefting, etc., and argues that this is so because Conjunction-Reduction is a copying and deletion, rather than a chopping, rule. However, Conjunction-Reduction is the only rule considered by Neubauer which operates across-the-board. Since the various chopping rules are also free to operate across-the-board, and since it would be exceedingly ad hoc to regard them as copying and deletion rules when operating across-the-board and as chopping ones otherwise, the inevitable conclusion is that chopping rules do not apply across-the-board to coordinate terms, but rather to the output of some copying and deletion rules. The only copying and deletion rule which applies to coordinations in English is Conjunction-Reduction, and Neubauer's position implicitly claims that (A4.6c) is derived from (A4.6a) via (A4.6b) rather than directly.

(A4.6) a. John likes potatoes and Mary hates potatoes.
b. John likes, and Mary hates, potatoes.
c. It's potatoes that John likes and Mary hates.

Such an analysis must be rejected for several reasons:

(a) The pauses indicated by commas in (A4.6b) are obligatory, while no such requirement holds for (A4.6c). Conceivably, one might claim that pause insertion is a later rule than Clefting, and that the former somehow depends on the moved element being to the right of its place of departure. But such a claim would dissolve in the face of (A4.7) in which the pauses are perfectly optional, although the moved element is to the right of its places of departure in surface structure.

(A4.7) What John likes and Mary hates is potatoes.

(b) Conjunction-Reduction can only touch initial and final elements of coordinate terms, but chopping rules can touch medial elements as well. Therefore, (A4.8a) would have no source, since (A4.8b) is ungrammatical and (A4.8c), while grammatical, arises through Topicalization rather than through Conjunction-Reduction.

(A4.8) a. It's John who I saw in Paris yesterday and you saw in London the day before yesterday.
b. *I saw in Paris yesterday, and you saw in London the day before yesterday, John.
c. John(,) I saw in Paris yesterday and you saw in London the day before yesterday.

(c) As noted above, across-the-board chopping is possible, with various degrees of acceptability, in non-coordinate structures. However, Conjunction-Reduction is completely out in such cases, as shown in (A4.9)-(A4.11).

(A4.9) a. It's Mary who John likes more than Bill hates.
b. *John likes, more than Bill hates, Mary.

(A4.10) a. It's John Mary told that she loves.
b. *Mary told, that she loves, John.

(A4.11) a. It's John who the undisputable effort to save suggests a hidden intention to destroy.
b. *The undisputable effort to save, suggests a hidden intention to destroy, John.

The only way of upholding Neubauer's position that I can see is to posit a copying and deletion rule, call it Mysterious-Reduction, which would derive (A4.6b) from (A4.6a) without requiring pause insertion, which would also derive structures like the b-sentences in (A4.8)-(A4.11), and which would moreover be obligatorily followed by some chopping rule applying to the constituent it has copied. I assume that the artificiality of such a proposal is sufficiently obvious for the matter not to be pursued any further.

BIBLIOGRAPHY

Akmajian, A. 1970. On Deriving Cleft Sentences from Pseudo-Cleft Sentences. *Linguistic Inquiry* 1.149-168.

Bach, E. 1968. Nouns and Noun Phrases, in Bach and Harms, eds., *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, 90-122.

Bach, E. 1969. Antipronominalization. University of Texas.

Bach, E. 1971. Questions. *Linguistic Inquiry* 2.153-166.

Baker, C. L. 1968. *Indirect Questions in English*. University of Illinois unpublished Ph.D. dissertation.

Baker, C. L., and M. K. Brame. 1972. 'Global rules': a rejoinder. *Language* 48.51-75.

Bever, T. G. 1970. The Cognitive Basis for Linguistic Structures, In Hayes, Ed., *Cognition and the Development of Language*. New York: Wiley, 277-360.

Bever, T. G. 1971. The Integrated Study of Language Behavior, in Mouton, Ed., *Psycholinguistics*. Logos Press, 158-210.

Bever, T. G., and D. T. Langendoen. 1970. The interaction of perception and grammar in linguistic change, in Stockwell and Macauley, eds., *Historical Linguistics in the Perspective of Transformational Theory*. Indiana University Press (1972).

Bresnan, J. 1970. On Complementizers: towards a syntactic theory of complement types. *Foundations of Language* 6.297-321.

Chapin, G. P., T. S. Smith, and A. A. Abrahamson. 1972. Two factors in perceptual segmentation of speech. *Journal of Verbal Learning and Verbal Behavior* 11.164-173.

Chomsky, N. A. 1964a. The logical basis of linguistic theory, in H. Lunt, ed., *Proceedings of the Ninth International Congress of Linguists*, Cambridge, Mass., 1962. The Hague: Mouton and Co.

Chomsky, N. A. 1964b. Current issues in linguistic theory, in Fodor and Katz, eds., *The Structure of Language*. Englewood Cliffs, N.J.: Prentice-Hall, 50-118.

Chomsky, N. A. 1969. Deep structure, surface structure and semantic interpretation, in Steinberg and Jakobovits, eds., *Semantics, An Interdisciplinary Reader in Philosophy, Linguistics and Psychology.* Cambridge, Cambridge University Press, 183-216.

Chomsky, N. A. 1970a. Remarks on nominalization, in Jacobs and Rosenbaum, eds., *Readings in English Transformational Grammar*. Watham, Mass.: Blaisdell-Ginn, 184-221.

Chomsky, N. A. 1970b. Empirical issues in the theory of transformational grammar. Indiana University Linguistics Club.

Chomsky, N. A. 1971. Conditions on Transformations, M.I.T.

Chomsky, N. A., and G. A. Miller. 1963. Introduction to the formal analysis of natural languages, in Luce, Buch, and Galenter, eds., *Handbook of Mathematical Psychology*, Vol. 2, New York: Wiley, 269-321.

Dean, J. 1967. Noun Phrase Complementation in English and German. M.I.T.

Dougherty, R. 1970. A grammar of coordinate conjoined structures: I. Language 46.850-898.

Drachman, G. 1970. Copying and order changing transformations in Greek. Working Papers in Linguistics No. 4, Ohio State University.

Elliott, D. 1971. The Grammar of Emotive and Exclamatory Predicates in English. Ph.D. dissertation, in Working Papers in Linguistics No. 8, The Ohio State University.

Emmonds, J. 1969a. Root and Structure Preserving Transformations. M.I.T. Ph.D. dissertation. Indiana University Linguistics Club.

Emmonds, J. 1969b. A structure-preserving constraint on NP movement transformations. Papers from the Fifth Regional Meeting of the Chicago Linguistic Society, 60-65.

Eriksen, C. W., and H. J. Johnson. 1964. Storage and decay characteristics of non-attended auditory stimuli. Journal of Experimental Psychology 68.423-434.

Fodor, J., and M. Garrett. 1966. Some reflections on competence and performance, in J. Lyons and Wales, eds., Psycholinguistics Papers. Edinburgh University Press, 132-185.

Geis, M. 1970. Adverbial Subordinate Clauses in English. M.I.T. unpublished Ph.D. dissertation.

Geis, M., and A. Zwicky. 1971. On invited inferences. Linguistic Inquiry 2.561-566.

Green, G. 1971. Notes on clefts and pseudo-clefts and other related matters. Studies in the Linguistic Sciences 1, University of Illinois.

Grice, H. P. 1968. The Logic of Conversation. Unpublished Manuscript.

Grinder, J. 1969. Conjunct Splitting in Samoan, Linguistic Notes from La Jolla II.

Grosu, A. 1970. On coreferentiality constraints and Equi-NP-Deletion in English. M. A. thesis, in Working Papers in Linguistics No. 7. Ohio State University.

Grosu, A. 1971. On perceptual and grammatical constraints, Papers from the Seventh Regional Meeting of the Chicago Linguistic Society, 416-427.

Grosu, A. 1972. Is the A-over-A principle an A-over-A principle? Papers from the Eighth Regional Meeting of the Chicago Linguistic Society.

Grosu, A. Forthcoming a. A note on implicatures, invited inferences and syntactic form. Die Sprache, 19.

Grosu, A. Forthcoming b. On the status of the so-called 'right roof constraint.' Language 49.

Guttman, N., and Julesz B. 1963. Lower limits of auditory periodicity analysis. Journal of the Acoustical Society of America 35.610.

Hankamer, J. 1971 . Constraints on Deletion in Syntax. Yale University unpublished Ph.D. dissertation.

Jackendoff, R. 1969. Some Rules of Semantic Interpretation for English. M.I.T. unpublished Ph.D. dissertation.

Koutsoudas, A. 1968. The A-over-A convention. Linguistics 46.11-20.

Koutsoudas, A., G. Sanders, and Nell. 1971. On the application of phonological rules. Indiana University Linguistics Club.

Lakoff, G. 1968. On pronouns and reference. Harvard University.

Lakoff, G. 1969. On derivational constraints. *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, 117-139.

Lakoff, G. 1970a. *Irregularity in Syntax*. New York: Holt, Rinehart, and Winston.

Lakoff, G. 1970b. Global rules. *Language* 46.627-639.

Lakoff, G. 1970c. Some thoughts on transderivational constraints. University of Michigan.

Lakoff, G. 1971. The role of deduction in grammar, in Fillmore and Langendoen, eds., *Studies in Linguistic Semantics*. New York: Holt, Rinehart and Winston, 63-72.

Lakoff, G. 1972. The arbitrary basis of transformational grammar. *Language* 48.76-87.

Lakoff, G., and S. Peters. 1969. Phrasal conjunction and symmetric predicates, in Reibel and Schane, eds., *Modern Studies in English, Readings in Transformational Grammar*. Englewood Cliffs, N.J.: Prentice-Hall, 113-142.

Lakoff, R. 1969. Some reasons why there can't be any *some-any* rule. *Language* 45.608-615.

Lakoff, R. 1971. And's, If's and But's about conjunction, in Fillmore and Langendoen, eds., *Studies in Linguistic Semantics*. New York: Holt, Rinehart and Winston, 115-150.

Lakoff, R. 1972. Review of *Progress in Linguistics: A Collection of Papers*, Bierwisch and Heidolph, eds., The Hague: Mouton, 1970. University of California, Berkeley.

Langacker, R. 1969. On pronominalization and the chain of command, in Reibel and Schane, eds., *Modern Studies in English, Readings in Transformational Grammar*. Englewood Cliffs, N. J.: Prentice-Hall, 160-186.

Lehiste, I. 1972. The Units of Speech Perception. *Working Papers in Linguistics* No. 12, Ohio State University.

Lehman, T. 1971. On the ordering of PARTICLE MOVEMENT, EXTRAPOSITION, QUESTION, and EXTRAPOSITION-FROM-NP. Indiana University Linguistics Club.

McCawley, J. 1971. Tense and time reference in English, in Fillmore and Langendoen, eds., *Studies in Linguistic Semantics*. New York: Holt, Rinehart and Winston, 97-114.

Mehler, J., and P. Carey. 1968. The interaction of veracity and syntax in the processing of sentences. *Perception and Psychophysics* 3.109-111.

Miller, G. A. 1962. Some psychological studies in grammar. *The American Psychologist* 17.748-762.

Neisser, U. 1967. *Cognitive Psychology.* New York: Appleton-Century-Crofts.

Neubauer, P. 1970. On the notion 'chopping rule', in *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, 400-407.

Nobel, B. 1971. Evidence. *Working Papers in Linguistics* No. 8, Ohio State University.

Peters, S., and E. Bach, 1968. Pseudo-cleft sentences. University of Texas.

Pollack, I. 1959. Message uncertainty and message reception. *Journal of the Acoustical Society of America* 31.1500-1508.

Postal, P. 1970a. *A Derivational Constraint on Pronominalization*. Yorktown Heights, N.Y.: I.B.M. Corporation Watson Research Center.

Postal, P. 1970b. On coreferential complement subject deletion. *Linguistic Inquiry* 1.439-500.

Postal, P. 1971. *Cross-over Phenomena*. New York: Holt, Rinehart and Winston.

Poutsma, H. 1929. *A Grammar of Late Modern English*. Part I, First Half, *The Composite Sentence* (2nd ed.) Groningen: P. Nordhoff.

Quang-Phuc-Dong. 1970. A note on conjoined noun phrases. *Journal of Philosophical Linguistics* 2.31-40.

Robinson, J. 1970. Dependency structures and transformational rules. *Language* 6.259-285.

Rosenbaum, P. S. 1967 . *The Grammar of English Predicate Complement Constructions*. Cambridge, Mass.: M.I.T. Press.

Ross, J. R. 1967. *Constraints on Variables in Syntax*. Unpublished M.I.T. Ph.D. dissertation.

Ross, J. R. 1969a. A proposed rule of tree-pruning, in Reibel and Schane, eds., *Modern Studies in English, Readings in Transformational Grammar*. Englewood Cliffs, N.J.: Prentice-Hall, 288-300.

Ross, J. R. 1969b. Guess who? *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, 252-286.

Ross, J. R. 1971. Act, in Baridson and Hartman, eds., *Semantics of Natural Languages*. Dordrecht, Holland: D. Reidel, 80-136.

Ross, J. R. 1972, The category hierarchy: Endstation Hauptwort, *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*.

Schmerling, S. F. 1972. Apparent counterexamples to the coordinate structure constraint: a canonical conspiracy. *Studies in the Linguistic Sciences* 2.1, University of Illinois

Shopen, T. 1972. Logical equivalence is not semantic equivalence, *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*, 340-350.

Stockwell, R., P. Schachter, and B. H. Partee. 1968. *Integration of Transformational Theories on English Syntax*. ESD-TR-68-419, U.S. Air Force.

Tai, J. H. 1969. *Coordination Reduction*. Indiana University unpublished Ph.D. dissertation.

Vendler, Z. 1961. Order of adjectives. *Transformations and Discourse Analysis Papers* 31. University of Pennsylvania.

Zwicky, A. M. 1969. Phonological constraints in syntactic description. *Papers in Linguistics* 1.411-463.